

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:20 ; Search time 40.2909 Seconds  
(without alignments)  
3570.253 Million cell updates/sec

Title: US-10-665-283-1  
Perfect score: 9903  
Sequence: 1 MALRGFCSDGSLDWDNV.....RSVAVAKAKPKFSIPDLSL 1927

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/6C\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7860	79.4	1531	1	US-08-463-092B-4
2	7860	79.4	1531	2	US-08-462-109A-4
3	7860	79.4	1531	2	US-08-460-907B-4
4	7860	79.4	1531	3	US-08-463-179A-4
5	7860	79.4	1531	3	US-08-461-384B-4
6	7860	79.4	1531	4	US-09-647-140B-19
7	7849	79.3	1531	1	US-08-141-893-2
8	7849	79.3	1531	1	US-08-463-092B-2
9	7849	79.3	1531	2	US-08-462-109A-2
10	7849	79.3	1531	2	US-08-460-907B-2
11	7849	79.3	1531	3	US-08-463-179A-2
12	7849	79.3	1531	3	US-08-461-384B-2
13	7849	79.3	1531	3	US-08-407-207A-2
14	7002.5	70.7	1528	1	US-08-463-092B-6
15	7002.5	70.7	1528	2	US-08-462-109A-6
16	7002.5	70.7	1528	2	US-08-460-907B-6
17	7002.5	70.7	1528	3	US-08-463-179A-6
18	7002.5	70.7	1528	3	US-08-461-384B-6
19	4487.5	45.3	1527	4	US-09-647-140B-6
20	4487.5	45.3	1530	4	US-09-647-140B-33
21	3391.5	34.2	1503	4	US-09-647-140B-8
22	3385.5	34.2	1503	4	US-09-792-616-3
23	3218.5	32.5	1498	4	US-09-792-616-9
24	2374.5	24.0	1621	3	US-08-972-927-3
25	2349.5	23.7	1622	3	US-08-972-927-6
26	2266.5	22.9	1325	4	US-09-647-140B-2
27	2195.5	22.2	1261	3	US-09-439-313-538

28	2195.5	22.2	1261	4	US-09-636-215-538
29	2195.5	22.2	1261	4	US-09-685-166A-538
30	2195.5	22.2	1261	4	US-09-679-426-538
31	2195.5	22.2	1261	4	US-09-759-143-538
32	2195.5	22.2	1261	4	US-09-651-236-538
33	2118	21.4	1581	3	US-08-726-320-3
34	2118	21.4	1581	3	US-09-208-716-3
35	2115	21.4	1228	3	US-09-439-313-537
36	2115	21.4	1228	4	US-09-636-215-537
37	2115	21.4	1228	4	US-09-685-166A-537
38	2115	21.4	1228	4	US-09-679-426-537
39	2115	21.4	1228	4	US-09-759-143-537
40	2115	21.4	1228	4	US-09-651-236-537
41	2082	21.0	1437	3	US-09-061-400-2
42	2082	21.0	1453	2	US-09-001-273-2
43	2082	21.0	1453	3	US-08-843-459A-2
44	2081	21.0	1437	4	US-10-162-012-38
45	2080	21.0	1437	4	US-09-647-140B-4

## ALIGNMENTS

RESULT 1  
US-08-463-092B-4  
; Sequence 4, Application US/08463092B  
; Patent No. 5766880  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
; STREET: Queen's University at Kingston  
; CITY: Kingston  
; STATE: Ontario  
; COUNTRY: CANADA  
; ZIP: K7L 3N6  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,092B  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steeg, Carol Miernicki  
; REGISTRATION NUMBER: 39,539  
; REFERENCE/DOCKET NUMBER: Q1546  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (613) 545-2342  
; TELEFAX: (613) 545-6853  
; INFORMATION FOR SEQ ID NO: 4:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLSCULE TYPE: protein
US-08-463-092B-4

Query Match 79.4%; Score 7860; DB 1; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCGADGSDPLWDMNVTWNTSNPDTKCFQNTVLVWPCFYLMACFPFYLILSRH 60
DQ 1 MALRGFCGADGSDPLWDMNVTWNTSNPDTKCFQNTVLVWPCFYLMACFPFYLILSRH 60
QY 61 DRGYIQTMLNKTALGFLWICWADLFSYFWERSRGIFLAPFLVPSLTLLGTTLLA 120
DQ 61 DRGYIQTMLNKTALGFLWICWADLFSYFWERSRGIFLAPFLVPSLTLLGTTLLA 120
QY 121 TFLIQLERRKGQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDFRDIIFYVYFS 180
DQ 121 TFLIQLERRKGQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDFRDIIFYVYFS 180
QY 181 LLLIQLVLSGSDRPLSETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
DQ 181 LLLIQLVLSGSDRPLSETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKOPAKQKESKVDANEEVEAL 300
DQ 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKOPAKQKESKVDANEEVEAL 300
QY 301 IVKSPQKWNPSLFVLYKTRGYPFLMSFPFKAHDLMMFSGPQILKLLIKFVNDTKAPD 360
DQ 301 IVKSPQKWNPSLFVLYKTRGYPFLMSFPFKAHDLMMFSGPQILKLLIKFVNDTKAPD 360
QY 361 WQGYFYTVLLFWTACQLVLHQPFIHFVSGMRKTAIVGAVRKALVITNSARKSSTV 420
DQ 361 WQGYFYTVLLFWTACQLVLHQPFIHFVSGMRKTAIVGAVRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRFMDLTYINMITSAPLOVTLALYLMLNGLSPSVLGVAVMLVMPVN 480
DQ 421 GEIVNLSVDAQRFMDLTYINMITSAPLOVTLALYLMLNGLSPSVLGVAVMLVMPVN 480
QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDLVLAIRBELKVLK 540
DQ 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDLVLAIRBELKVLK 540
QY 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVTTDENNILDQAQTAFAVSLALFNILRPLNLP 600
DQ 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVTTDENNILDQAQTAFAVSLALFNILRPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLGSHELEPDSIERPPVKDGGGNSITVRNATFTWARSDDPT 660
DQ 601 MVISSIVQASVSLKRLRIFLGSHELEPDSIERPPVKDGGGNSITVRNATFTWARSDDPT 660
QY 661 LNGITFSIPEGALVAVVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720
DQ 661 LNGITFSIPEGALVAVVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720
QY 721 SLRENILFGCCLEBPYRSVITQACALLPDLBILSGDRTEIGEGVNLGGQKQKVSILAR 780
DQ 721 SLRENILFGCCLEBPYRSVITQACALLPDLBILSGDRTEIGEGVNLGGQKQKVSILAR 780
QY 781 AVYSNADILYFDDPLSAVDHVGKHIENVIIPGKMLKNKTRILVTHSMSYLPQVDVILV 840
DQ 781 AVYSNADILYFDDPLSAVDHVGKHIENVIIPGKMLKNKTRILVTHSMSYLPQVDVILV 840
QY 841 MSGGKISMGYSQELLARDGAFBLRTYASTEQDAEENGVTGVSQPGKEAKQOMENG 900
DQ 841 MSGGKISMGYSQELLARDGAFBLRTYASTEQDAEENGVTGVSQPGKEAKQOMENG 900
QY 901 LVTDGAGKQOLQSSSSSYSGDLSRHNSAEELQKAEAKKEETWKLMEADKAQTQGVKL 960
DQ 901 LVTDGAGKQOLQSSSSSYSGDLSRHNSAEELQKAEAKKEETWKLMEADKAQTQGVKL 960

```

```

RESULT 2
US-08-462-109A-4
; Sequence 4, Application US/08462109A
; Patent No. 5882875
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,109A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340

```

```

; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Decont, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCES/DOCKET NUMBER: PQI-002CP4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-462-109A-4

Query Match          79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGCAGSDPLMDNVTWNTSNPDKFQNTLVVWPCFYLMACFPFYFLYLSRH 60
DB 1 MALRGCAGSDPLMDNVTWNTSNPDKFQNTLVVWPCFYLMACFPFYFLYLSRH 60
QY 61 DRGYTQMTPLNKTALGFLWICWADLFYSFWERSRGIFLAPVFLSPTLGLITLLA 120
DB 61 DRGYTQMTPLNKTALGFLWICWADLFYSFWERSRGIFLAPVFLSPTLGLITLLA 120
QY 121 TFLIQLERRKGVSQSGIMLTFWLVALVICALALIRSKIMTALKEDAQVDLFRDITYVYFS 180
DB 121 TFLIQLERRKGVSQSGIMLTFWLVALVICALALIRSKIMTALKEDAQVDLFRDITYVYFS 180
QY 181 LLLIQLVSCFSDRSPLETHDNPCESSASFLSRITFWITGLIVRGVQPLEGSD 240
DB 181 LLLIQLVSCFSDRSPLETHDNPCESSASFLSRITFWITGLIVRGVQPLEGSD 240
QY 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPVVYSSKDPKQKESKVDANEVEAL 300
DB 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPVVYSSKDPKQKESKVDANEVEAL 300
QY 301 IVKSPQKWNPSLFKVLKTFPGYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
DB 301 IVKSPQKWNPSLFKVLKTFPGYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
QY 361 WQGYFTVLLFVTAQLTLVHLQYFHCIFVSGMRKTAIVAGVYRKALVITNSARKSTV 420
DB 361 WQGYFTVLLFVTAQLTLVHLQYFHCIFVSGMRKTAIVAGVYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVTLALYLWNLGSPSLAGVAVMLVFPVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVTLALYLWNLGSPSLAGVAVMLVFPVN 480
QY 481 AVAMKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKYLAWELAFKDKVLAIROBELKVLK 540
DB 481 AVAMKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKYLAWELAFKDKVLAIROBELKVLK 540
QY 541 KSAYLSAVGTFTWCTPPELVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600
DB 541 KSAYLSAVGTFTWCTPPELVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600
QY 601 WVISSIVQASVSLKRLRFLSHEELEPDSIERPPKDGCGTNSITVRNATFTWARSDDPT 660
DB 601 WVISSIVQASVSLKRLRFLSHEELEPDSIERPPKDGCGTNSITVRNATFTWARSDDPT 660
QY 661 LNGITFTSIPEGALVAVVGQVGCKSKSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIOND 720
DB 661 LNGITFTSIPEGALVAVVGQVGCKSKSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIOND 720
QY 721 SLRENILFGCQLEPYRSVIOACALLPDLPLEIPSGDRTEIGEKGWNLGGQKQVSLAR 780

```

```

DB 721 SLRENILFGCQLEPYRSVIOACALLPDLPLEIPSGDRTEIGEKGWNLGGQKQVSLAR 780
QY 781 AVYSNADIYLPDDPLSADVAHVGHKHFENVIGPKGMLKNKTRILVTHSMYLPQVDVLIIV 840
DB 781 AVYSNADIYLPDDPLSADVAHVGHKHFENVIGPKGMLKNKTRILVTHSMYLPQVDVLIIV 840
QY 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEQQDAEENGVTGVSFGPKGKAKOMENGM 900
DB 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEQQDAEENGVTGVSFGPKGKAKOMENGM 900
QY 901 LVTDGAGLQORQLSSSSSYSGDISRHNSHNSAELOKABAKKEETWKLMEADKAQTGVKVL 960
DB 901 LVTDGAGLQORQLSSSSSYSGDISRHNSHNSAELOKABAKKEETWKLMEADKAQTGVKVL 960
QY 961 SVYWDYMKAIGLFISFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTKVRLSVYG 1020
DB 961 SVYWDYMKAIGLFISFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTKVRLSVYG 1020
QY 1021 ALGISQGIAGVFGYSMAVISGIGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSEL 1080
DB 1021 ALGISQGIAGVFGYSMAVISGIGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSEL 1080
QY 1081 DTVDSMIPDEVTKMPMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140
DB 1081 DTVDSMIPDEVTKMPMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140
QY 1141 KRLSVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIIVANRWLA 1200
QY 1201 VRLECVGNCIVLFAALFAVIRSRHSLSAGLVLSYSLSQVTTYLNWLVMSSEMETNIVA 1260
DB 1201 VRLECVGNCIVLFAALFAVIRSRHSLSAGLVLSYSLSQVTTYLNWLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETKEAPWQIQTAPPSSWPQGRVFRNTCLRYREDLDFVLRHINTVINGG 1320
DB 1261 VERLKEYSETKEAPWQIQTAPPSSWPQGRVFRNTCLRYREDLDFVLRHINTVINGG 1320
QY 1321 EKVGIIVGRTGAGKSLTLGLFRINESARGEIIDIINIAKIGLHDLRFKTIIPDPPVLF 1380
DB 1321 EKVGIIVGRTGAGKSLTLGLFRINESARGEIIDIINIAKIGLHDLRFKTIIPDPPVLF 1380
QY 1381 SGSLRMNLDPFSQYSDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440
DB 1381 SGSLRMNLDPFSQYSDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440
QY 1441 ARALLRKTIIIVLDEATAAVIDETDDLLIQSTIRQFEDCTVLTIAHRLNTIMDYTRVIVL 1500
DB 1441 ARALLRKTIIIVLDEATAAVIDETDDLLIQSTIRQFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGETOEYGAPESSDILQORGLFYSMAKDAGLV 1531
DB 1501 DKGETOEYGAPESSDILQORGLFYSMAKDAGLV 1531

```

```

RESULT 3
US-08-460-907B-4
; Sequence 4, Application US/08460907B
; Patent No. 5891724
; GENERAL INFORMATION:
; APPLICANT: Deeley, Roger G.
; APPLICANT: Cole, Susan P.C.
; TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
; TITLE OF INVENTION: RESISTANCE ON A CELL
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6

```

```

/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: ASCII text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/460,907B
/ FILING DATE: 05-JUN-1995
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/966,923
/ FILING DATE: 27-OCT-1992
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/029,340
/ FILING DATE: 8-MAR-1993
/ CLASSIFICATION: 424
/ APPLICATION DATA:
/ APPLICATION NUMBER: 08/141,893
/ FILING DATE: 26-OCT-1993
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/407,207
/ FILING DATE: 20-MAR-1995
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Steeg, Carol Miernicki
/ REGISTRATION NUMBER: 39,539
/ REFERENCE/DOCKET NUMBER: Q1551
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (613) 545-2342
/ TELEFAX: (613) 545-6853
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1531 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-460-907B-4

Query Match 79.4%; Score 7860; DB 2; Length 1531;
Besc Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCADGSDPLMDNNTWNTSNPDTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60
DB 1 MALRGFCADGSDPLMDNNTWNTSNPDTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60
QY 61 DRGIQMTPLNKTALGFLWIVCWADLFYSFWERSRGI FLAPVFLVSP TLLGITTLLA 120
DB 61 DRGIQMTPLNKTALGFLWIVCWADLFYSFWERSRGI FLAPVFLVSP TLLGITTLLA 120
QY 121 TFLQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLPEDITFYVYFS 180
DB 121 TFLQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLPEDITFYVYFS 180
QY 181 LLLIQLVLSCFSDRSPFSETHDPNCPESASFLSRITFWITGLIVRGYRPLEGSD 240
DB 181 LLLIQLVLSCFSDRSPFSETHDPNCPESASFLSRITFWITGLIVRGYRPLEGSD 240
QY 241 LWSLNKETSQVVPVLVKNWKECAKTRKQPKVYVSSKDPAPQKSSKVDANEVEAL 300
DB 241 LWSLNKETSQVVPVLVKNWKECAKTRKQPKVYVSSKDPAPQKSSKVDANEVEAL 300
QY 301 IVKSPQKEWNPFLFKVLYKTFPGPYFLMSFFPKAHDLMWFSGPOILKLLIKFVNDTKAPD 360
DB 301 IVKSPQKEWNPFLFKVLYKTFPGPYFLMSFFPKAHDLMWFSGPOILKLLIKFVNDTKAPD 360
QY 361 WQGYFYTVLLFVTAQLOTLVLHQYFHCFFVSGMRKIKTAVIGAVYRKALVITNSARKSSTV 420
DB 361 WQGYFYTVLLFVTAQLOTLVLHQYFHCFFVSGMRKIKTAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLLNLNGLPSVLGAVVWMLVMPVN 480

```



Db 1501 DKGEIQEYCAPSDLLQORGLFYMAKDAGLV 1531

RESULT 4

US-08-463-179A-4  
 ; Sequence 4, Application US/08463179A  
 ; Patent No. 6001563  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cole, Susan P.C.  
 ; APPLICANT: Deeley, Roger G.  
 ; TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD  
 ; STREET: 60 State Street, suite 510  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02109

COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/463,179A  
 ; FILING DATE:  
 ; CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/966,923  
 ; FILING DATE: 27-OCT-1992  
 ; APPLICATION NUMBER: 08/029,340  
 ; FILING DATE: 8-MAR-1993  
 ; APPLICATION NUMBER: 08/141,893  
 ; FILING DATE: 26-OCT-1993  
 ; APPLICATION NUMBER: 08/407,207  
 ; FILING DATE: 20-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: DeConti, Giulio A. Jr.  
 ; REGISTRATION NUMBER: 31,503  
 ; REFERENCE/DOCKET NUMBER: PQI-002CP8  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5941  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1531 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-463-179A-4

Query Match 79.4%; Score 7860; DB 3; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSDGSDPLMDNVTWNTSNPDFTKCFQNTLVWVPCFYLMACFPFYLYLSRH 60  
 |||||  
 Db 1 MALRGFCSDGSDPLMDNVTWNTSNPDFTKCFQNTLVWVPCFYLMACFPFYLYLSRH 60  
 |||||

QY 61 DRGYIOMTPLNKTALGELLMIWCVADLFYFWSRSGIFLAPVFLVSPDLLGTTLLA 120  
 |||||  
 Db 61 DRGYIOMTPLNKTALGELLMIWCVADLFYFWSRSGIFLAPVFLVSPDLLGTTLLA 120  
 |||||

QY 121 TELIOLERRKGVOSGIMLTFWLVALCALATLRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 |||||  
 Db 121 TELIOLERRKGVOSGIMLTFWLVALCALATLRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 |||||

QY 181 LLLIOLVLSCFSDRSPFLSETIHDNPPCPSSASFLSRITFWITGLIVRGVROPLEGSD 240  
 |||||  
 Db 181 LLLIOLVLSCFSDRSPFLSETIHDNPPCPSSASFLSRITFWITGLIVRGVROPLEGSD 240  
 |||||

QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPKVYSSKDPAPQKSSKVDANEVEAL 300  
 |||||

Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPKVYSSKDPAPQKSSKVDANEVEAL 300  
 |||||

QY 301 IVKSPKEWNPDLFKVLYKTFPGPYFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 |||||

Db 301 IVKSPKEWNPDLFKVLYKTFPGPYFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 |||||

QY 361 WQGYFTVLLFVTACLOTFLVHQYPHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
 |||||

Db 361 WQGYFTVLLFVTACLOTFLVHQYPHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
 |||||

QY 421 GEIVNLSMSVDAQRFMDLATYINMIWSAPLOVTLALYLWLNLPSPVLAGVAVMLMVPYN 480  
 |||||

Db 421 GEIVNLSMSVDAQRFMDLATYINMIWSAPLOVTLALYLWLNLPSPVLAGVAVMLMVPYN 480  
 |||||

QY 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLKYLAWELAFKDKVLATROEBELKVLK 540  
 |||||

Db 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLKYLAWELAFKDKVLATROEBELKVLK 540  
 |||||

QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
 |||||

Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
 |||||

QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDCGGTNSITVRNATFTWARSDDPT 660  
 |||||

Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDCGGTNSITVRNATFTWARSDDPT 660  
 |||||

QY 661 LMGITFISPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIWD 720  
 |||||

Db 661 LMGITFISPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIWD 720  
 |||||

QY 721 SLRENILFCQLEBPYRSVIAQACALLPDLLEIPSGDRTEIGEKGKVNLSGGQKQVSLAR 780  
 |||||

Db 721 SLRENILFCQLEBPYRSVIAQACALLPDLLEIPSGDRTEIGEKGKVNLSGGQKQVSLAR 780  
 |||||

QY 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 |||||

Db 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 |||||

QY 841 MSGGKISEMGYSQELLARDGAFABFLRTVASTEQEOADAEENGVTGSGPGKEAKWENGM 900  
 |||||

Db 841 MSGGKISEMGYSQELLARDGAFABFLRTVASTEQEOADAEENGVTGSGPGKEAKWENGM 900  
 |||||

QY 901 LVTDGKQQLQRLSSSSSSSGDISRHHNSTAELOKAEAKKEETWKLMEADKAQTCQVKL 960  
 |||||

Db 901 LVTDGKQQLQRLSSSSSSSGDISRHHNSTAELOKAEAKKEETWKLMEADKAQTCQVKL 960  
 |||||

QY 961 SVYWDYKKAIGLFIISFLSIFLFCMCHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 |||||

Db 961 SVYWDYKKAIGLFIISFLSIFLFCMCHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 |||||

QY 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTSGNLVNRFSKEL 1080  
 |||||

Db 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTSGNLVNRFSKEL 1080  
 |||||

QY 1081 DTVDSMIEVIKMFMSGLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 |||||

Db 1081 DTVDSMIEVIKMFMSGLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 |||||

QY 1141 KRLSVSRSPYSHPNETLLGVSVIRAPBEOERFIHQSDLKVDENQKAYIPSVIANRWLA 1200  
 |||||

Db 1141 KRLSVSRSPYSHPNETLLGVSVIRAPBEOERFIHQSDLKVDENQKAYIPSVIANRWLA 1200  
 |||||

QY 1201 VRLECVGNCIVLFAALFAVIRSHSISAGLVLSYSISLQVTTYLNLWLRMSSEMETNIVA 1260  
 |||||

Db 1201 VRLECVGNCIVLFAALFAVIRSHSISAGLVLSYSISLQVTTYLNLWLRMSSEMETNIVA 1260  
 |||||

QY 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNRYCLARYREDLDFVLRHINTINGG 1320  
 |||||

Db 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNRYCLARYREDLDFVLRHINTINGG 1320  
 |||||

QY 1321 EKVGIWGTGAGKSLTLGLPRINESAGEEIIIDGINIAKIGLHDLRPFKITIIPDQVULF 1380  
 |||||

Db 1321 EKVGIVGRTGACKSSLTGLFRINESAGEIITDGINIAKIGHDLRFKTIIPQDPVLF 1380  
 Qy 1381 SGLRNLDPFQSYDEEYVWTSLELAHLKDFVSALPDKLDHCAEGENLSVGQRLVCL 1440  
 Db 1381 SGLRNLDPFQSYDEEYVWTSLELAHLKDFVSALPDKLDHCAEGENLSVGQRLVCL 1440  
 Qy 1441 ARALLRKTILVDEATAVDDLETDLLIOSTIQTQPEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTILVDEATAVDDLETDLLIOSTIQTQPEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Qy 1501 DKGEIQEYCAPSDLLQORGLFYSMAKDAGLV 1531  
 Db 1501 DKGEIQEYCAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 5

US-08-461-384B-4  
 ; Sequence 4, Application US/08461384B  
 ; Patent No. 6025473  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cole, Susan P.C.  
 ; APPLICANT: Deeley, Roger G.  
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
 ; STREET: Queen's University at Kingston  
 ; CITY: Kingston  
 ; STATE: Ontario  
 ; COUNTRY: CANADA  
 ; ZIP: K7L 3N6  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/461.384B  
 ; FILING DATE: 05-JUN-95  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/966,923  
 ; FILING DATE: 27-OCT-1992  
 ; APPLICATION NUMBER: 08/029,340  
 ; FILING DATE: 8-MAR-1993  
 ; APPLICATION NUMBER: 08/141,893  
 ; FILING DATE: 26-OCT-1993  
 ; APPLICATION NUMBER: 08/407,207  
 ; FILING DATE: 20-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Steeg, Carol Miernicki  
 ; REGISTRATION NUMBER: 39,539  
 ; REFERENCE/DOCKET NUMBER: Q1547  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (613) 545-2342  
 ; TELEFAX: (613) 545-6853  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1531 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-461-384B-4

Query Match 79.4%; Score 7860; DB 3; Length 1531;  
 Best Local Similarity 100.08; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MALRGFCADGSDPLWDNNVTWNTSNPDFTKCFQNTLVWVPCFYLWACFPFFLYLSRH 60  
 Db 1 MALRGFCADGSDPLWDNNVTWNTSNPDFTKCFQNTLVWVPCFYLWACFPFFLYLSRH 60  
 Qy 61 DRGVIQMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPILLGITLLA 120

Db 61 DRGVIQMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPILLGITLLA 120  
 Qy 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDFRDITPYVYVFS 180  
 Db 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDFRDITPYVYVFS 180  
 Qy 181 LLLLIQVLSCFSDRSPFLSETHDPNCPSESSASPLSRITFWWITGLVIRGVROPLEGSD 240  
 Db 181 LLLLIQVLSCFSDRSPFLSETHDPNCPSESSASPLSRITFWWITGLVIRGVROPLEGSD 240  
 Qy 241 LWSLNKEDTSEQVVPVLVKNWKKECAKTRKQPVKVYVYSSKDPAPQPKSSKVDANEVEAL 300  
 Db 241 LWSLNKEDTSEQVVPVLVKNWKKECAKTRKQPVKVYVYSSKDPAPQPKSSKVDANEVEAL 300  
 Qy 301 IVKSPQKEWNPFLFKVLYKTFPGPYFLMSFFPKAHDLMFSGPQILKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKEWNPFLFKVLYKTFPGPYFLMSFFPKAHDLMFSGPQILKLLIKFVNDTKAPD 360  
 Qy 361 WQGYFYTVLLFVTAQCLQTLVHLQYFHCIVFSGMRKTAIVIGAVYRKALAVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTAQCLQTLVHLQYFHCIVFSGMRKTAIVIGAVYRKALAVITNSARKSSTV 420  
 Qy 421 GEIVNLSVDAQREFMDLATYINMIWSAPLQVILALYLLNLNLGSPSVLAGVAVMLVMPVN 480  
 Db 421 GEIVNLSVDAQREFMDLATYINMIWSAPLQVILALYLLNLNLGSPSVLAGVAVMLVMPVN 480  
 Qy 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAPKDKVLAIROBELKVLK 540  
 Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAPKDKVLAIROBELKVLK 540  
 Qy 541 KSAVLSAVGTFTWCTPFLVALCTFAVVVTIDENNILDAOTAFVLSLAFNLIRPLNLTP 600  
 Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVVVTIDENNILDAOTAFVLSLAFNLIRPLNLTP 600  
 Qy 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWARSPPPT 660  
 Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWARSPPPT 660  
 Qy 661 LNGTHTESTPEGALVAVVGQCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWTQND 720  
 Db 661 LNGTHTESTPEGALVAVVGQCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWTQND 720  
 Qy 721 SLRENILFGQLEPYPYRSVIQACALLPDLLEILPSGDRTEIGEKNVLSGQKQKQVSLAR 780  
 Db 721 SLRENILFGQLEPYPYRSVIQACALLPDLLEILPSGDRTEIGEKNVLSGQKQKQVSLAR 780  
 Qy 781 AVYSNADIYLPDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840  
 Db 781 AVYSNADIYLPDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840  
 Qy 841 MSGGKISEMGSYQELLARDGAFEFRTYASTEQDAEENGVTGVSQPGKEAKOMENGM 900  
 Db 841 MSGGKISEMGSYQELLARDGAFEFRTYASTEQDAEENGVTGVSQPGKEAKOMENGM 900  
 Qy 901 LVTDSAGKQLQRLQSSSSSYSGDLSRHNSSTAEIQLKAEKEETWKLMEADKATQGVKL 960  
 Db 901 LVTDSAGKQLQRLQSSSSSYSGDLSRHNSSTAEIQLKAEKEETWKLMEADKATQGVKL 960  
 Qy 961 SVYWDYMKALGLFISFLSIFLFCNHNVSALASNWLSTLWDDPIVNGTQHTKVLRSYVG 1020  
 Db 961 SVYWDYMKALGLFISFLSIFLFCNHNVSALASNWLSTLWDDPIVNGTQHTKVLRSYVG 1020  
 Qy 1021 ALGISQIGIAVFGYSMAVSIIGIILASRCLHVDLLHSILRSWSPFPFRTSGNLVNRFSKEL 1080  
 Db 1021 ALGISQIGIAVFGYSMAVSIIGIILASRCLHVDLLHSILRSWSPFPFRTSGNLVNRFSKEL 1080  
 Qy 1081 DTVDSMIPEVKMPMGSFLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140  
 Db 1081 DTVDSMIPEVKMPMGSFLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140  
 Qy 1141 KRLESVRSRSPVYSHFNETHLIGSVIRAFERQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
 Db 1141 KRLESVRSRSPVYSHFNETHLIGSVIRAFERQERFIHQSDLKVDENQKAYPSIVANRWLA 1200

Qy	1201	VRLECVGNCIVLFAALPAVISRHSLSAGLVGLSVSYSLQVTYYINLWLRMSSEMTNIVA	1260
Db	1201	VRLECVGNCIVLFAALPAVISRHSLSAGLVGLSVSYSLQVTYYINLWLRMSSEMTNIVA	1260
Qy	1261	VERLKEYSETEKAPWOIQETAPPSSWPQGRVEFRNYCLRYREDLDVLRLHNVITNGG	1320
Db	1261	VERLKEYSETEKAPWOIQETAPPSSWPQGRVEFRNYCLRYREDLDVLRLHNVITNGG	1320
Qy	1321	EKVGIVGRTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380
Db	1321	EKVGIVGRTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380
Qy	1381	SGSLRMNLDPFSQYSDEEWTSLLELAHLKDFVSALPKDLHECAEGGENISVGQRQLVCL	1440
Db	1381	SGSLRMNLDPFSQYSDEEWTSLLELAHLKDFVSALPKDLHECAEGGENISVGQRQLVCL	1440
Qy	1441	ARALLRKTILVLDEATAAVDLETDDLIQSTIRTPQEDCTVLTAHRLNTIMDYTRVIVL	1500
Db	1441	ARALLRKTILVLDEATAAVDLETDDLIQSTIRTPQEDCTVLTAHRLNTIMDYTRVIVL	1500
Qy	1501	DKGEIOYGAPSDLLQORGLFYSGMAKDAGLV	1531
Db	1501	DKGEIOYGAPSDLLQORGLFYSGMAKDAGLV	1531

181	Db	LLLLQLVLSCFSDRSPLFSETHDNPDPRESSASFJLSRITFWITGLIVRGVRQPLEGSD	240
241	Qy	LWSLNKEDTSEQVVPVLVKNWKKCEKATRKQPVKVVYSSKDAQPKESSKVDANEVEAL	300
241	Db		
241	Db	LWSLNKEDTSEQVVPVLVKNWKKCEKATRKQPVKVVYSSKDAQPKESSKVDANEVEAL	300
301	Qy	IVKSPQKWNPSLFKVLKYKTFGPYFLMSPPFKAIHDLMMPSGQILKLLIKFVNDTKAPD	360
301	Db		
301	Db	IVKSPQKWNPSLFKVLKYKTFGPYFLMSPPFKAIHDLMMPSGQILKLLIKFVNDTKAPD	360
361	Qy	WQGYFYTVLLFVTAQLQTLVLHGVFHCIPVSGMIRIKTAVIGAVYRKALVITNSARKSSTV	420
361	Db		
361	Db	WQGYFYTVLLFVTAQLQTLVLHGVFHCIPVSGMIRIKTAVIGAVYRKALVITNSARKSSTV	420
421	Qy	GEIYNLMSVDAQRFMDLATYINNIWSAPLQVILALYLLMLNLGSPVLGAVVMVLMPVN	480
421	Db		
421	Db	GEIYNLMSVDAQRFMDLATYINNIWSAPLQVILALYLLMLNLGSPVLGAVVMVLMPVN	480
481	Qy	AVMAWTKTQVVAHMSKONRIIKLMEILNGIKVLKLYAWELAFKQKVLAIROEELKVLK	540
481	Db		
481	Db	AVMAWTKTQVVAHMSKONRIIKLMEILNGIKVLKLYAWELAFKQKVLAIROEELKVLK	540
541	Qy	KSAYLSAVGFTTWVCTPFLVALCTFAVYVTIDENNILDAQTAFCVSLALFNILRFPNLNLP	600
541	Db		
541	Db	KSAYLSAVGFTTWVCTPFLVALCTFAVYVTIDENNILDAQTAFCVSLALFNILRFPNLNLP	600
601	Qy	MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKDGCGTNSITVRNATFTTWARS DPPT	660
601	Db		
601	Db	MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKDGCGTNSITVRNATFTTWARS DPPT	660
661	Qy	LNGITPSPGALVAVVQVGCCKSSLLSALLAEMDKVEGHVAIKGSVATVPOQAWIQND	720
661	Db		
661	Db	LNGITPSPGALVAVVQVGCCKSSLLSALLAEMDKVEGHVAIKGSVATVPOQAWIQND	720
721	Qy	SLRENILFGQLEPPYRYSYQACALLPDEIILPSGDRTEIGBKGNLSGGQKQRVSLAR	780
721	Db		
721	Db	SLRENILFGQLEPPYRYSYQACALLPDEIILPSGDRTEIGBKGNLSGGQKQRVSLAR	780
781	Qy	AVYSNADIYLFDDPLSAVDARVGHKIFENVIGPKGMLKNKTRILVTHSHMSYLPQVDVIIV	840
781	Db		
781	Db	AVYSNADIYLFDDPLSAVDARVGHKIFENVIGPKGMLKNKTRILVTHSHMSYLPQVDVIIV	840
841	Qy	MSGKISEMGSYQELLARDGAPAEFLRTYASTSQEQDAEBENGVTGSGPQKEAKQWENG	900
841	Db		
841	Db	MSGKISEMGSYQELLARDGAPAEFLRTYASTSQEQDAEBENGVTGSGPQKEAKQWENG	900
901	Qy	LVTDTSAGKQIQRLSSSSSYSGDISRHNSHSTABLOKAEAKKEETWKLMEADKAQTQGVKL	960
901	Db		
901	Db	LVTDTSAGKQIQRLSSSSSYSGDISRHNSHSTABLOKAEAKKEETWKLMEADKAQTQGVKL	960
961	Qy	SVYWDYMKALGLFISLFIPLFCMNVHSALASNYWLSLWTDDEPIVNGTOHHTKVRLSVYG	1020
961	Db		
961	Db	SVYWDYMKALGLFISLFIPLFCMNVHSALASNYWLSLWTDDEPIVNGTOHHTKVRLSVYG	1020
1021	Qy	ALGISQIGIAVFGYSMAVSIQGLILASRCLHVDLLHSILRSPMSPFERTPSGNLVNRFSKEL	1080
1021	Db		
1021	Db	ALGISQIGIAVFGYSMAVSIQGLILASRCLHVDLLHSILRSPMSPFERTPSGNLVNRFSKEL	1080
1081	Qy	DTVDSMTPEVIKNPMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVYVASSRQL	1140
1081	Db		
1081	Db	DTVDSMTPEVIKNPMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVYVASSRQL	1140
1141	Qy	KRLSVSRSPVYGHFNFTLLGVSVIRAFEBQERFIHQSDLKVDENOKAYYPSIVANRWLA	1200
1141	Db		
1141	Db	KRLSVSRSPVYGHFNFTLLGVSVIRAFEBQERFIHQSDLKVDENOKAYYPSIVANRWLA	1200
1201	Qy	VRLECVGNCIVLFAALPAVISHRSHLSAGLVGLSVSYSLQVTTVTLNLVLRMSSEMETNIVA	1260
1201	Db		
1201	Db	VRLECVGNCIVLFAALPAVISHRSHLSAGLVGLSVSYSLQVTTVTLNLVLRMSSEMETNIVA	1260
1261	Qy	VERLUKEYSETEKAPWQIQETAPSSWPQVGRVEFRNYCLRYREDLDFVLRLHNVINGG	1320
1261	Db		
1261	Db	VERLUKEYSETEKAPWQIQETAPSSWPQVGRVEFRNYCLRYREDLDFVLRLHNVINGG	1320

QY 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGHDLRFKXITIIPODPVL 1380  
DB 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGHDLRFKXITIIPODPVL 1380  
QY 1381 SGLSRNLDPFQSDDEEYVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL 1440  
DB 1381 SGLSRNLDPFQSDDEEYVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL 1440  
QY 1441 ABALLRKTILVLDENATAVDLETDLLIOSTTRTQFEDCTVLTIAHRLNTIMDYTRVIL 1500  
DB 1441 ABALLRKTILVLDENATAVDLETDLLIOSTTRTQFEDCTVLTIAHRLNTIMDYTRVIL 1500  
QY 1501 DKGEIQEYCAPSDLLQORGLFYSMADAGLV 1531  
DB 1501 DKGEIQEYCAPSDLLQORGLFYSMADAGLV 1531

## RESULT 7

US-08-141-893-2  
Sequence 2, Application US/08141893  
Patent No. 5489519  
GENERAL INFORMATION:  
APPLICANT: Cole, Susan P.C.  
APPLICANT: Deeley, Roger G.  
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEIN  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSES: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/141,893  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/966,923; 08/029,340  
FILING DATE: 27-OCT-1992; 8-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Deconti, Giulio A. Jr.  
REGISTRATION NUMBER: 31,503  
REFERENCE/DOCKET NUMBER: PQI-002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5149  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1531 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-141-893-2

Query Match 79.3%; Score 7849; DB 1; Length 1531;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSDGSDPLDWDNVNTWNTSNPDTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60  
DB 1 MALRGFCSDGSDPLDWDNVNTWNTSNPDTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60  
QY 61 DRGVIQMTPLNKTALGFLWVWVWADLFFYSFWRSGIFLAPVFLVSPFLIGITLLA 120  
DB 61 DRGVIQMTPLNKTALGFLWVWVWADLFFYSFWRSGIFLAPVFLVSPFLIGITLLA 120

QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAIALIRSKINTALKEDAQVDLFRDITPVYVFS 180  
DB 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAIALIRSKINTALKEDAQVDLFRDITPVYVFS 180  
QY 181 LLLIQLVLSGSDRSPLESETHDNPCESSASFLSRITFWWITGLIVRGVQRLGSD 240  
DB 181 LLLIQLVLSGSDRSPLESETHDNPCESSASFLSRITFWWITGLIVRGVQRLGSD 240  
QY 241 LWSLNKEDTSQVVPVLVKWKKCECAKTRKQPVKVYSSKDPAPQKESKVDANEVEAL 300  
DB 241 LWSLNKEDTSQVVPVLVKWKKCECAKTRKQPVKVYSSKDPAPQKESKVDANEVEAL 300  
QY 301 IVKSPQKEMNLSFKVLYKTFPGPYLMSFFPKAIDHLMFSGPQILKLIKFVNDTKAPD 360  
DB 301 IVKSPQKEMNLSFKVLYKTFPGPYLMSFFPKAIDHLMFSGPQILKLIKFVNDTKAPD 360  
QY 361 WQGYFYTVLLFVTACLOTVLHQYFHCFSVGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
DB 361 WQGYFYTVLLFVTACLOTVLHQYFHCFSVGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLNLNLGSPSVLAGVAMVLMVNVN 480  
DB 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLNLNLGSPSVLAGVAMVLMVNVN 480  
QY 481 AVMAKTKTYOVAMHMKDNRIKLMNETLNGIKVLKYAWELAFKDKVLATROBELKVLK 540  
DB 481 AVMAKTKTYOVAMHMKDNRIKLMNETLNGIKVLKYAWELAFKDKVLATROBELKVLK 540  
QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRPLNLP 600  
DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRPLNLP 600  
QY 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKCGGTSITVRNATWTWASDPT 660  
DB 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKCGGTSITVRNATWTWASDPT 660  
QY 661 LNGITFSIPEGALVAVVGVCGCKSLLSALLASMDKVEGHVAKGVSAYVVPQAWIOND 720  
DB 661 LNGITFSIPEGALVAVVGVCGCKSLLSALLASMDKVEGHVAKGVSAYVVPQAWIOND 720  
QY 721 SLRENILFGCOLEBPYYSRVIQACALLPDLLEILPSGDRTEIGEGVNLSSGQKQVSLAR 780  
DB 721 SLRENILFGCOLEBPYYSRVIQACALLPDLLEILPSGDRTEIGEGVNLSSGQKQVSLAR 780  
QY 781 AVYSNADIYLFDDPLSAVDAHVGKHIPENVIGPKMLKNKTRILVTHSMSVLPQVDVIL 840  
DB 781 AVYSNADIYLFDDPLSAVDAHVGKHIPENVIGPKMLKNKTRILVTHSMSVLPQVDVIL 840  
QY 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEQDAEENGVTGVSQPGKEAKQEMNGM 900  
DB 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEQDAEENGVTGVSQPGKEAKQEMNGM 900  
QY 901 LVTDGAKQLQRLSSSSSYSGDISRHHNSTAEILOKAEKKEETWKLMEADKQAGVVKL 960  
DB 901 LVTDGAKQLQRLSSSSSYSGDISRHHNSTAEILOKAEKKEETWKLMEADKQAGVVKL 960  
QY 961 SVYWDYMKALGLFISFLSIFLEWCNHYVSALSNVLSLWTDPIVNGTQHTKVLVSYG 1020  
DB 961 SVYWDYMKALGLFISFLSIFLEWCNHYVSALSNVLSLWTDPIVNGTQHTKVLVSYG 1020  
QY 1021 ALGISQGIAGVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL 1080  
DB 1021 ALGISQGIAGVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL 1080  
QY 1081 DTVDSDMTPEVIMKPMGSLFNVIAGACIVILLATPIAIIIPPLGLIYFFVORFYVASSRQL 1140  
DB 1081 DTVDSDMTPEVIMKPMGSLFNVIAGACIVILLATPIAIIIPPLGLIYFFVORFYVASSRQL 1140  
QY 1141 KRLESVSRSPYSHFNETLLGVSVIRAFERQERFIHQSLDKVDENQKAYVPSIVANRWLA 1200  
DB 1141 KRLESVSRSPYSHFNETLLGVSVIRAFERQERFIHQSLDKVDENQKAYVPSIVANRWLA 1200  
QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGSYSLSQVTTVYLNWLVNRMSSMETNIVA 1260

Db 1201 VLELCVGNICVILFAALFAVISRHSLSAGLVGSYSLSQVTTYNLWVMSSEMETNIVA 1260  
Qy 1261 VERLKEYSETEKAPWQIQTETAPPSSWQVGRVFRNRYCLARYREDLDFVLRHINTINGG 1320  
Db 1261 VERLKEYSETEKAPWQIQTETAPPSSWQVGRVFRNRYCLARYREDLDFVLRHINTINGG 1320  
Qy 1321 EKVGVIGRTGAGKSLITLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
Db 1321 EKVGVIGRTGAGKSLITLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
Qy 1381 SGLSRMLDPPFQSYDEEVTWTSLELAHLKDFVSALPDKLDHECARGGENLSVGQRLVCL 1440  
Db 1381 SGLSRMLDPPFQSYDEEVTWTSLELAHLKDFVSALPDKLDHECARGGENLSVGQRLVCL 1440  
Qy 1441 ABALLRKTKILVLDATAVDETDLTQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
Db 1441 ABALLRKTKILVLDATAVDETDLTQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
Qy 1501 DKGEIQEYCAPSDLLQQRGLFYSMADAGLV 1531  
Db 1501 DKGEIQEYCAPSDLLQQRGLFYSMADAGLV 1531

RESULT 8

US-08-463-092B-2  
Sequence 2, Application US/08463092B  
Patent No. 5766880  
GENERAL INFORMATION:  
APPLICANT: Cole, Susan P.C.  
APPLICANT: Deeleey, Roger G.  
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING  
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
STREET: Queen's University at Kingston  
CITY: Kingston  
STATE: Ontario  
COUNTRY: CANADA  
ZIP: K7L 3N6  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,092B  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/966,923  
FILING DATE: 27-OCT-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/029,340  
FILING DATE: 8-MAR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/141,893  
FILING DATE: 26-OCT-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/407,207  
FILING DATE: 20-MAR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Steeg, Carol Miernicki  
REGISTRATION NUMBER: 39,539  
REFERENCE/DOCKET NUMBER: Q1546  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (613) 545-2342  
TELEFAX: (613) 545-6853

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1531 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-463-092B-2  
Query Match 79.3%; Score 7849; DB 1; Length 1531;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 MALRGFCASDGSDFLMDWNVTWNTSNPDTKCFQNTVLVWVPCFVLMACFPFFYLYLSRH 60  
Db 1 MALRGFCASDGSDFLMDWNVTWNTSNPDTKCFQNTVLVWVPCFVLMACFPFFYLYLSRH 60  
Qy 61 DRGIQMTPLNKTALGFLMIWCVADLFFYSFWEBSRGI FLAPVFLVSPITLLGTTLLA 120  
Db 61 DRGIQMTPLNKTALGFLMIWCVADLFFYSFWEBSRGI FLAPVFLVSPITLLGTTLLA 120  
Qy 121 TFLIQLERRKGVQSSGIMLTFLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
Db 121 TFLIQLERRKGVQSSGIMLTFLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
Qy 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPSSASPLSRITFWITGLIVRGYRQPLEGSD 240  
Db 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPSSASPLSRITFWITGLIVRGYRQPLEGSD 240  
Qy 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKOPKVYVSSKDPAPKESKVDANEVEAL 300  
Db 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKOPKVYVSSKDPAPKESKVDANEVEAL 300  
Qy 301 IVKSPQKEWNPFLFKVLYKTGPFYFLMSFFFKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360  
Db 301 IVKSPQKEWNPFLFKVLYKTGPFYFLMSFFFKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360  
Qy 361 WQGYFTVLLFVTACLTQLVLHQYPHICFVSGMRIKTAVIGAVYRKALVITNSARKSTV 420  
Db 361 WQGYFTVLLFVTACLTQLVLHQYPHICFVSGMRIKTAVIGAVYRKALVITNSARKSTV 420  
Qy 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVTLALYLALNGLSPSLAGVAVMLVMPVN 480  
Db 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVTLALYLALNGLSPSLAGVAVMLVMPVN 480  
Qy 481 AVMAKTKTYQVAHMKSKDNRIKLANEILNGIKVLKYAWELAFKDKVLAIQEEELKVLK 540  
Db 481 AVMAKTKTYQVAHMKSKDNRIKLANEILNGIKVLKYAWELAFKDKVLAIQEEELKVLK 540  
Qy 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLIIP 600  
Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLIIP 600  
Qy 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660  
Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660  
Qy 661 LMGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720  
Db 661 LMGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720  
Qy 721 SLRENILFGCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
Db 721 SLRENILFGCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
Qy 781 AVYSNADYILFDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILYTHSNLYLPQVDVILV 840  
Db 781 AVYSNADYILFDDPLSAVDHVGKHI FENVIGPKGMLKNKTRILYTHSNLYLPQVDVILV 840  
Qy 841 MSGGKISEMGSYQELLARDGAFABFLRTVASTEQDQDAEENGVTGVSFGPKAKOMENGM 900  
Db 841 MSGGKISEMGSYQELLARDGAFABFLRTVASTEQDQDAEENGVTGVSFGPKAKOMENGM 900  
Qy 901 LVTDSAGKQLQRLSSSSSSSGSDISRHHNSTAELQKAEAKBETWKLMBADKAQTQGVKL 960

Db	901	LVTDSAGQIQORQLSSSSSYGDISRHNSSTAELQKAEAKCBETWKLMEADKQTCQVKL	960
Qy	961	SVYWDYMKALGLFTISLFIFLFCMNHVSALASNYWLSLWTDDEPVTNGTOBHTKVRLSVYG	1020
Db	961	SVYWDYMKALGLFTISLFIFLFCMNHVSALASNYWLSLWTDDEPVTNGTOBHTKVRLSVYG	1020
Qy	1021	ALGISQIAPVFGYSMAVSIIGILASRCLHVDLLHSILRSPMSPFERTPSGNLVNRFSKEL	1080
Db	1021	ALGISQIAPVFGYSMAVSIIGILASRCLHVDLLHSILRSPMSPFERTPSGNLVNRFSKEL	1080
Qy	1081	DTVDSMIPEVIKMFMSGLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVVASRQL	1140
Db	1081	DTVDSMIPEVIKMFMSGLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVVASRQL	1140
Qy	1141	KRLSVRSRSPVYGHFNETLLGVSVIRAFESQERFIHQSDLKVDENOKAYYPSIVANRWLA	1200
Db	1141	KRLSVRSRSPVYGHFNETLLGVSVIRAFESQERFIHQSDLKVDENOKAYYPSIVANRWLA	1200
Qy	1201	VRLECVGNCIVLFAALFAVLSRHSLSAGLVGLSVSYSLQVTTYVNLVLRMSSSEMETNIVA	1260
Db	1201	VRLECVGNCIVLFAALFAVLSRHSLSAGLVGLSVSYSLQVTTYVNLVLRMSSSEMETNIVA	1260
Qy	1261	VERLKEYSETEKAPWIOQETAPSSWPQGRVEFRNYCLRYREDLDVLRHNVITNGG	1320
Db	1261	VERLKEYSETEKAPWIOQETRPSSWPQGRVEFRNYCLRYREDLDVLRHNVITNGG	1320
Qy	1321	EKVGIVGRTGAGKSSLTGLFRINESAEGBIIIDGINIAKIGHDLRFKTIITIPQDPVLV	1380
Db	1321	EKVGIVGRTGAGKSSLTGLFRINESAEGBIIIDGINIAKIGHDLRFKTIITIPQDPVLV	1380
Qy	1381	SGSLRWNLDPFSQYSDEEVTSLLELAHLKDFVSALPDKLDHECAEGENLSVGORQLVCL	1440
Db	1381	SGSLRWNLDPFSQYSDEEVTSLLELAHLKDFVSALPDKLDHECAEGENLSVGORQLVCL	1440
Qy	1441	ARALLRKTILVLDEATAAVDLTDDLIQSTIRTFQEDCTVLTIARLNTIMDYTRVIVL	1500
Db	1441	ARALLRKTILVLDEATAAVDLTDDLIQSTIRTFQEDCTVLTIARLNTIMDYTRVIVL	1500
Qy	1501	DKGEI0EYGAPSDLLQOQGLFYFMAKDAGIV	1531
Db	1501	DKGEI0EYGAPSDLLQOQGLFYFMAKDAGIV	1531

## RESULT 9

US-08-462-109A-2  
Sequence 2, Application US/08462109A  
Patent No. 5882875  
GENERAL INFORMATION:  
APPLICANT: Cole, Susan P.C.  
APPLICANT: Deeley, Roger G.  
TITLE OF INVENTION: METHODS FOR IDENTIFYING  
TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, suite 510  
City: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462,109A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/966,923  
FILING DATE: 27-OCT-1992

```

QY 721 SIRENIFGQLEEPYRSVIOACALLPDLPLPSGDRTEIGEKGWNLGGQKQSVLAR 780
Db 721 SIRENIFGQLEEPYRSVIOACALLPDLPLPSGDRTEIGEKGWNLGGQKQSVLAR 780
QY 781 AYSNADIYLPDPLSAVDHVGKHFENVIKPKMLKNKTRILVTHSMYSYLPQVDVIV 840
Db 781 AYSNADIYLPDPLSAVDHVGKHFENVIKPKMLKNKTRILVTHSMYSYLPQVDVIV 840
QY 841 MSGGKISEMGYSQELLARDAFAELRTVASTEQODAEENGVTGSGPGKEAKOMENGM 900
Db 841 MSGGKISEMGYSQELLARDAFAELRTVASTEQODAEENGVTGSGPGKEAKOMENGM 900
QY 901 LVTDSAGKOLORQSSSSSYSGDISRHNNSTAELOKAEAKETWKLMEADKAQGTQVKL 960
Db 901 LVTDSAGKOLORQSSSSSYSGDISRHNNSTAELOKAEAKETWKLMEADKAQGTQVKL 960
QY 961 SYVDYMKAGLFIPLSTIFLPMCHVHSALASNYLSLWTDPIVNGTQHTKVLRSYVG 1020
Db 961 SYVDYMKAGLFIPLSTIFLPMCHVHSALASNYLSLWTDPIVNGTQHTKVLRSYVG 1020
QY 1021 ALGISQGIAGVGSYMAVSGGILASRCLHVDLLHSLRSPMSFFERTPSGNLVNRFSKEL 1080
Db 1021 ALGISQGIAGVGSYMAVSGGILASRCLHVDLLHSLRSPMSFFERTPSGNLVNRFSKEL 1080
QY 1081 DTVDMSIPEVIRKPMGSLPNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140
Db 1081 DTVDMSIPEVIRKPMGSLPNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140
QY 1141 KLESVSRSPVSHNETLLGVSVIRAFEEQERFIHQSDLKVDENKQKAYPSIVANRWLA 1200
Db 1141 KLESVSRSPVSHNETLLGVSVIRAFEEQERFIHQSDLKVDENKQKAYPSIVANRWLA 1200
QY 1201 VRLCEVGNICIVLFAALFAVISRHSLSAGLVGLSVSLQVTTYLNLVMSSEMETNIVA 1260
Db 1201 VRLCEVGNICIVLFAALFAVISRHSLSAGLVGLSVSLQVTTYLNLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPWQIETAPPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKAPWQIETAPPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
QY 1321 EKVGIVRTGAGKSSLTGLFRINSAEGEIIIDGINAKIGHDLRPFKTIIPQDPVLF 1380
Db 1321 EKVGIVRTGAGKSSLTGLFRINSAEGEIIIDGINAKIGHDLRPFKTIIPQDPVLF 1380
QY 1381 SGSLRNLDPPFSQYSDDEEVTLSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440
Db 1381 SGSLRNLDPPFSQYSDDEEVTLSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440
QY 1441 ARALLRKTILVLDDEATRAVDLETDLDLIQSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 ARALLRKTILVLDDEATRAVDLETDLDLIQSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIQEYCAPSDLLQOQGLFYSMADAGLV 1531
Db 1501 DKGEIQEYCAPSDLLQOQGLFYSMADAGLV 1531

```

```

RESULT 10
US-08-460-907B-2
; Sequence 2, Application US/08460907B
; Patent No. 5891724
; GENERAL INFORMATION:
; APPLICANT: Deeley, Roger G.
; APPLICANT: Cole, Susan P.C.
; TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
; TITLE OF INVENTION: RESISTANCE ON A CELL
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSES: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA

```

```

; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,907B
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1551
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-460-907B-2

```

```

Query Match 79.3%; Score 7849; DB 2; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCADGSDPLMDWNTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60
Db 1 MALRGFCADGSDPLMDWNTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60
QY 61 DRGYIQTMTPLNKTKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPITLLGITTLA 120
Db 61 DRGYIQTMTPLNKTKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPITLLGITTLA 120
QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVICALAILRSKIMTALKEDAQVDLFRUITFYVYPS 180
Db 121 TFLIQLERRKGVQSSGIMLTFWLVALVICALAILRSKIMTALKEDAQVDLFRUITFYVYPS 180
QY 181 LLLIQLVLSCFSDRSPLESETIHDNPNCPSSASFLSRITFWITGLIVRGVROPLEGSD 240
Db 181 LLLIQLVLSCFSDRSPLESETIHDNPNCPSSASFLSRITFWITGLIVRGVROPLEGSD 240
QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPKVYVSSKDPAPQPKSSKVDANBEVEAL 300
Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPKVYVSSKDPAPQPKSSKVDANBEVEAL 300
QY 301 IVKSPQKEWNPFLFKVLYKTFGPVFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEWNPFLFKVLYKTFGPVFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
QY 361 WQGYFYTVLLFVTAQLOTLVLHOYFHCIVSGMRKTKAVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGYFYTVLLFVTAQLOTLVLHOYFHCIVSGMRKTKAVIGAVYRKALVITNSARKSSTV 420

```



QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLLWNLGSPVLGAVVWLVMPVN 480  
DB |||||  
QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLLWNLGSPVLGAVVWLVMPVN 480  
DB |||||  
QY 481 AVMAKTKTYQVAHMKSKDNRIKLNWEILNGIKVLYAWELAFKDKVLAIRQEBELKVK 540  
DB |||||  
QY 481 AVMAKTKTYQVAHMKSKDNRIKLNWEILNGIKVLYAWELAFKDKVLAIRQEBELKVK 540  
DB |||||  
QY 541 KSAYLSAVGTFTWCTPPLVALCTFAVYVTDENNILDAQTAFAVSLALFNILRFPNLILP 600  
DB |||||  
QY 541 KSAYLSAVGTFTWCTPPLVALCTFAVYVTDENNILDAQTAFAVSLALFNILRFPNLILP 600  
DB |||||  
QY 601 MYISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWASDPT 660  
DB |||||  
QY 601 MYISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWASDPT 660  
DB |||||  
QY 661 LNGITFSIPEGALVAVGVGCGKSLLSALLAEMDKVEGHVAIRKSVAYVPPQAWIQND 720  
DB |||||  
QY 661 LNGITFSIPEGALVAVGVGCGKSLLSALLAEMDKVEGHVAIRKSVAYVPPQAWIQND 720  
DB |||||  
QY 721 SURENILFGCOLPEEYPSVIOACALLPDLLEILPSGDRTEIGEGKGNLSGGOKORVSLAR 780  
DB |||||  
QY 721 SURENILFGCOLPEEYPSVIOACALLPDLLEILPSGDRTEIGEGKGNLSGGOKORVSLAR 780  
DB |||||  
QY 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIGPKGMLKXKTRILVTHSMSYLPQVDVIV 840  
DB |||||  
QY 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIGPKGMLKXKTRILVTHSMSYLPQVDVIV 840  
DB |||||  
QY 841 MSGGKISEMGYSQELLARDGAPAEFLRTYVASTEQDAEENGVTGSGFGKEAKOMENGM 900  
DB |||||  
QY 841 MSGGKISEMGYSQELLARDGAPAEFLRTYVASTEQDAEENGVTGSGFGKEAKOMENGM 900  
DB |||||  
QY 901 LVTDAGLQORQLSSSSYSYSGDISRHNSSTAELQAKAEKBEWKLMEADKAQGOVKL 960  
DB |||||  
QY 901 LVTDAGLQORQLSSSSYSYSGDISRHNSSTAELQAKAEKBEWKLMEADKAQGOVKL 960  
DB |||||  
QY 961 SYVDYMKAIGLFISFLIFPMCNHVSALASNYWLSLWTDPIVNGTQEHTKVLRSYVG 1020  
DB |||||  
QY 961 SYVDYMKAIGLFISFLIFPMCNHVSALASNYWLSLWTDPIVNGTQEHTKVLRSYVG 1020  
DB |||||  
QY 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
DB |||||  
QY 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
DB |||||  
QY 1081 DTVDSMIEVIVKMGSLFNVIACIVILLATPIAAIIPPLGLIYFFVQRYFVASSROL 1140  
DB |||||  
QY 1081 DTVDSMIEVIVKMGSLFNVIACIVILLATPIAAIIPPLGLIYFFVQRYFVASSROL 1140  
DB |||||  
QY 1141 KELESVSRSPVYSHFNETLLGVSIVIRAFEEQERFIHQSDLKVDENQKAYIPSIVANRWLA 1200  
DB |||||  
QY 1201 VLECVGNCIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNLWVRMSSEMETNIVA 1260  
DB |||||  
QY 1201 VLECVGNCIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNLWVRMSSEMETNIVA 1260  
DB |||||  
QY 1261 VERLKEYSETEKEAPWQOETAPPSPQVGVVERFNYCLRYREDLDFVLRHINVTINGG 1320  
DB |||||  
QY 1261 VERLKEYSETEKEAPWQOETAPPSPQVGVVERFNYCLRYREDLDFVLRHINVTINGG 1320  
DB |||||  
QY 1321 EVKGVGRTGAGKSLTLGLFRINSEAGEIIDIIGNIAKIGHDLRPFKTIIPQDPVLF 1380  
DB |||||  
QY 1321 EVKGVGRTGAGKSLTLGLFRINSEAGEIIDIIGNIAKIGHDLRPFKTIIPQDPVLF 1380  
DB |||||  
QY 1381 SGSLRMNLDPFQSYDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
DB |||||  
QY 1381 SGSLRMNLDPFQSYDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
DB |||||  
QY 1441 ABALLRKTKILVLDATAVDLETDLLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVILV 1500  
DB |||||  
QY 1441 ABALLRKTKILVLDATAVDLETDLLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVILV 1500  
DB |||||  
QY 1501 DKGEIQEYGAPSDLLQORGLFYSMADAGLV 1531

DB 1501 DKGEIQEYGAPSDLLQORGLFYSMADAGLV 1531

## RESULT 11

US-08-463-179A-2  
; Sequence 2, Application US/08463179A  
; Patent No. 6001563  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463.179A  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeConti, Giulio A. Jr.  
; REGISTRATION NUMBER: 31,503  
; REFERENCE/POCKET NUMBER: PQ1-002CF8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1531 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-463-179A-2

Query Match 79.3%; Score 7849; DB 3; Length 1531;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSDGSDPLNDWNVNTWNTSNDDFTKCFQNTVLVWVPCFYLWACFPFFFLYLSRH 60  
DB 1 MALRGFCSDGSDPLNDWNVNTWNTSNDDFTKCFQNTVLVWVPCFYLWACFPFFFLYLSRH 60  
QY 61 DRGIQMTPLNKTALGFLLIWVCWADLFYSFWSRSGIFLAPVFLVSPDLLGITLLA 120  
DB 61 DRGIQMTPLNKTALGFLLIWVCWADLFYSFWSRSGIFLAPVFLVSPDLLGITLLA 120  
QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
DB 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
QY 181 LLLQLVLSCFSDRSPLFSETHIDPNCPSSASFLSRITFWITGLIVRGVROPLEGSD 240  
DB 181 LLLQLVLSCFSDRSPLFSETHIDPNCPSSASFLSRITFWITGLIVRGVROPLEGSD 240



QY 241 LMSLNKEDTSEQVVPVLVQNMKKCEAKTRKQPVVYYSKDPAPKESKVDANEVEAL 300  
 DB 241 LMSLNKEDTSEQVVPVLVQNMKKCEAKTRKQPVVYYSKDPAPKESKVDANEVEAL 300  
 QY 301 IVKSPQKWNPSLKVLYKTEGYPYELMSFFFAIHDLMFSGPOTILKLLIKFVNDTKAPD 360  
 DB 301 IVKSPQKWNPSLKVLYKTEGYPYELMSFFFAIHDLMFSGPOTILKLLIKFVNDTKAPD 360  
 QY 361 MQGYFYTVLLFVYACIQLTLVHLQYFHI CFVSGMRKTAIVGAVYRKALVITNSARKSSTV 420  
 DB 361 MQGYFYTVLLFVYACIQLTLVHLQYFHI CFVSGMRKTAIVGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINNIWSAPLOVTLALYLLNLNLPSPVLAGVAVMLVMPVN 480  
 DB 421 GEIVNLSVDAQRFMDLATYINNIWSAPLOVTLALYLLNLNLPSPVLAGVAVMLVMPVN 480  
 QY 481 AVAMKTKTYQVAHMKSKDNRIKLANEILNGIKVLYAWELAPKDKVLAIRQBELKVLK 540  
 DB 481 AVAMKTKTYQVAHMKSKDNRIKLANEILNGIKVLYAWELAPKDKVLAIRQBELKVLK 540  
 QY 541 KSAYLSAVGTFTWVCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
 DB 541 KSAYLSAVGTFTWVCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERPVDGGGTSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERPVDGGGTSITVRNATFTWARSDDPT 660  
 QY 661 LNGITFSIPEGALVAVGVGGKSLASALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
 DB 661 LNGITFSIPEGALVAVGVGGKSLASALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
 QY 721 SURENILFGCQLEEPYYSVIOACALLPDLEILPSGDRTEIGEGKVNLSGGOKQVSLAR 780  
 DB 721 SURENILFGCQLEEPYYSVIOACALLPDLEILPSGDRTEIGEGKVNLSGGOKQVSLAR 780  
 QY 781 AVYSNADILYFDPLSADVAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 DB 781 AVYSNADILYFDPLSADVAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 QY 841 MSGGKISEMGSYOELLARDGAEFLRTYASTEQEODAEENGVTGSGGKEAKOMENGM 900  
 DB 841 MSGGKISEMGSYOELLARDGAEFLRTYASTEQEODAEENGVTGSGGKEAKOMENGM 900  
 QY 901 LVTDGAGKQLOQLSSSSYSYSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQGTQVKL 960  
 DB 901 LVTDGAGKQLOQLSSSSYSYSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQGTQVKL 960  
 QY 961 SVYDYMKAIGLFIISFLIFPMCNHVSALASNYWLSLWTDPIVNGTQHTKVRLSVYG 1020  
 DB 961 SVYDYMKAIGLFIISFLIFPMCNHVSALASNYWLSLWTDPIVNGTQHTKVRLSVYG 1020  
 QY 1021 ALGISQGTAVFGYSMAVSTGGTILASRCLHVDLLHSILRSPMSFFERTPSGNLWNRFSKEL 1080  
 DB 1021 ALGISQGTAVFGYSMAVSTGGTILASRCLHVDLLHSILRSPMSFFERTPSGNLWNRFSKEL 1080  
 QY 1081 DTVDSDMIEVKNFMGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVFQFYVASSRQL 1140  
 DB 1081 DTVDSDMIEVKNFMGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVFQFYVASSRQL 1140  
 QY 1141 KRLESVRSVPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
 DB 1141 KRLESVRSVPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISHSLSAGLVGSYSLOVTTYLNLVMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVISHSLSAGLVGSYSLOVTTYLNLVMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKAPQOETAPPSSWPQGVGFERNCLRYREDLDFVLRHINVTINGG 1320  
 DB 1261 VERLKEYSETEKAPQOETAPPSSWPQGVGFERNCLRYREDLDFVLRHINVTINGG 1320  
 QY 1321 EKVIGVGRGAGKSSLTGLFRINESABEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380

DB 1321 EKVIGVGRGAGKSSLTGLFRINESABEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SSGSRMNLDPFSQYSDEBWTSLSLAHLKDFVSALPDKLDHECAEGGNSVGOQLVCL 1440  
 DB 1381 SSGSRMNLDPFSQYSDEBWTSLSLAHLKDFVSALPDKLDHECAEGGNSVGOQLVCL 1440  
 QY 1441 ARALLRKTILVLDEATAAVALDLEDDLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ARALLRKTILVLDEATAAVALDLEDDLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGIOEYGAESDILQORGLFYSMAKDAGLV 1531  
 DB 1501 DKGIOEYGAESDILQORGLFYSMAKDAGLV 1531

RESULT 12  
 US-08-461-384B-2  
 ; Sequence 2, Application US/08461384B  
 ; Patent No. 6025473  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cole, Susan P.C.  
 ; APPLICANT: Deeley, Roger G.  
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
 ; STREET: Queen's University at Kingston  
 ; CITY: Kingston  
 ; STATE: Ontario  
 ; COUNTRY: CANADA  
 ; ZIP: K7L 3N6  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/461,384B  
 ; FILING DATE: 05-JUN-95  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/966,923  
 ; FILING DATE: 27-OCT-1992  
 ; APPLICATION NUMBER: 08/029,340  
 ; FILING DATE: 8-MAR-1993  
 ; APPLICATION NUMBER: 08/141,893  
 ; FILING DATE: 26-OCT-1993  
 ; APPLICATION NUMBER: 08/407,207  
 ; FILING DATE: 20-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Steeg, Carol Miernicki  
 ; REGISTRATION NUMBER: 39,539  
 ; REFERENCE/DOCKET NUMBER: Q1547  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (613) 545-2342  
 ; TELEFAX: (613) 545-6853  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1531 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-461-384B-2

Query Match 79.3%; Score 7849; DB 3; Length 1531;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MALRGFCSADGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFYLYLSRH 60  
 DB 1 MALRGFCSADGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFYLYLSRH 60  
 QY 61 DRGIQMTPLNKTALGFLWVCWADLFYSFWSRSRGI FLAPVFLVSPDLLGITLLA 120

Db 61 DRGYIQMTPLNKTALGFLMIWCVADLFYSFWERSRGIFLAPVFLVSPLLGITTLLA 120  
Qy 121 TFLIQLERRKGQSSGIMLTFWLVALVACALAILRSKIMTALKEDAQVDFLDRDITFYVYFS 180  
Db 121 TFLIQLERRKGQSSGIMLTFWLVALVACALAILRSKIMTALKEDAQVDFLDRDITFYVYFS 180  
Qy 181 LLLIQVLSCPSDRSPLRSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
Db 181 LLLIQVLSCPSDRSPLRSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
Qy 241 LWSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVYSSKDPAPQKBSKVDANEEVEAL 300  
Db 241 LWSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVYSSKDPAPQKBSKVDANEEVEAL 300  
Qy 301 IVKSPQKEWNPSPFLKVLVYKTPGYPFLMGFFPKAIDHLMFSGPQILKLLIKFVNDTKAPD 360  
Db 301 IVKSPQKEWNPSPFLKVLVYKTPGYPFLMGFFPKAIDHLMFSGPQILKLLIKFVNDTKAPD 360  
Qy 361 WQGYFYTVLLFWTACLOTILVHOYHICFVSGMRITKAVIGAVYRKALVITNSARKSSTV 420  
Db 361 WQGYFYTVLLFWTACLOTILVHOYHICFVSGMRITKAVIGAVYRKALVITNSARKSSTV 420  
Qy 421 GEIVNLSVDAORFMDLATYINMWSAPLOVILALYLWNLGSPVLGAVVAVMLVMPVN 480  
Db 421 GEIVNLSVDAORFMDLATYINMWSAPLOVILALYLWNLGSPVLGAVVAVMLVMPVN 480  
Qy 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
Qy 541 KSAYLSAVGTFTWCTPFLVACTFAVVTIDENNILDAQTAFAVSLAFNLIRPLNLPL 600  
Db 541 KSAYLSAVGTFTWCTPFLVACTFAVVTIDENNILDAQTAFAVSLAFNLIRPLNLPL 600  
Qy 601 WYISSIVQASVSLKELRIFLSHEELEPDSIERPVPKDGGSITSITVRNATFTWARSDDPT 660  
Db 601 WYISSIVQASVSLKELRIFLSHEELEPDSIERPVPKDGGSITSITVRNATFTWARSDDPT 660  
Qy 661 LINGITFSPEGALVAVGVQGGKSLLSALLAENDKVEGHVAIKGSVAYVPPQAWIQND 720  
Db 661 LINGITFSPEGALVAVGVQGGKSLLSALLAENDKVEGHVAIKGSVAYVPPQAWIQND 720  
Qy 721 SURENILFCQLEBPYRSVIOACALLPDLEILPSGDRTEIGEKVNLSSGGQKQSVLSAR 780  
Db 721 SURENILFCQLEBPYRSVIOACALLPDLEILPSGDRTEIGEKVNLSSGGQKQSVLSAR 780  
Qy 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
Db 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
Qy 841 MSGGKISEMGVQELLARDGAFELRTYASTEQDAEENGVTGVSQKQKQKQWNGM 900  
Db 841 MSGGKISEMGVQELLARDGAFELRTYASTEQDAEENGVTGVSQKQKQKQWNGM 900  
Qy 901 LVTDSAGKQLOQLSSSSSYSDISRHNSAELOKAEAKKEETWKLMEADKAQOGVKL 960  
Db 901 LVTDSAGKQLOQLSSSSSYSDISRHNSAELOKAEAKKEETWKLMEADKAQOGVKL 960  
Qy 961 SVYWDYMKAIGLIFGLSIFLPMCNHVSALASNYWLSLWTDTPVNGTQHTKVLRSVYG 1020  
Db 961 SVYWDYMKAIGLIFGLSIFLPMCNHVSALASNYWLSLWTDTPVNGTQHTKVLRSVYG 1020  
Qy 1021 ALGISOGIAVFGYMAVSTGGTILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEL 1080  
Db 1021 ALGISOGIAVFGYMAVSTGGTILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEL 1080  
Qy 1081 DTVDMSIPEVKMFNGSLFNVICACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140  
Db 1081 DTVDMSIPEVKMFNGSLFNVICACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140  
Qy 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFERQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200

Db 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFERQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200  
Qy 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVYSLOVTTYNLWLVNRMSSSEMTNIVA 1260  
Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVYSLOVTTYNLWLVNRMSSSEMTNIVA 1260  
Qy 1261 VERLKEYSETEKAPQIQETAPPSSWPQVGRVYEFYNYCLRYREDLDFVLRHINVTINGG 1320  
Db 1261 VERLKEYSETEKAPQIQETAPPSSWPQVGRVYEFYNYCLRYREDLDFVLRHINVTINGG 1320  
Qy 1321 EKVGIVGRTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
Db 1321 EKVGIVGRTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
Qy 1381 SGSLRMLNDPFSQYSDDEEVTSLLEHLKDPVSGALPKLDHECAEGGENLSVGORQLVCL 1440  
Db 1381 SGSLRMLNDPFSQYSDDEEVTSLLEHLKDPVSGALPKLDHECAEGGENLSVGORQLVCL 1440  
Qy 1441 ARALLRRTKILVLDEATAAVDLETDLLIQSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Db 1441 ARALLRRTKILVLDEATAAVDLETDLLIQSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Qy 1501 DKGBIQYAGAPSDLLQORGLFYSMADAGLV 1531  
Db 1501 DKGBIQYAGAPSDLLQORGLFYSMADAGLV 1531

## RESULT 13

US-08-407-207A-2

; Sequence 2, Application US/08407207A

; Patent No. 6063621

; GENERAL INFORMATION:

; APPLICANT: Deeley, Roger G.

; APPLICANT: Cole, Susan P.C.

; TITLE OF INVENTION: ANTIBODIES TO A MULTIDRUG RESISTANCE PROTEIN

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESS: PARTEQ RESEARCH &amp; DEVELOPMENT INNOVATIONS

; STREET: Queen's University at Kingston

; CITY: Kingston

; STATE: Ontario

; COUNTRY: CANADA

; ZIP: K7L 3N6

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/407,207A

; FILING DATE: 20-MAR-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/966,923

; FILING DATE: 27-OCT-1992

; APPLICATION NUMBER: 08/029,340

; FILING DATE: 8-MAR-1993

; APPLICATION NUMBER: 08/141,893

; FILING DATE: 26-OCT-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Steeg, Carol Miernicki

; REGISTRATION NUMBER: 39,539

; REFERENCE/DOCKET NUMBER: Q1512

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (613) 545-2342

; TELEFAX: (613) 545-6853

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1531 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-407-207A-2

Query Match		79.3%;	Score 7849;	DB 3;	Length 1531;	
Best Local Similarity		99.9%;	Pred. No. 0;			
Matches 1529;		Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;	
QY	1	MALRGCSDGSDPLWDMVNTWNTSNPDTKCFONTVLYVWPCFYLWACFPFYFLYLSRH	60			
DB	1	MALRGCSDGSDPLWDMVNTWNTSNPDTKCFONTVLYVWPCFYLWACFPFYFLYLSRH	60			
QY	61	DRGYQMTPLNKTALGFLLLWVWADLFYFWERSRGIFLAPVFLVSPPTLLGTTLLA	120			
DB	61	DRGYQMTPLNKTALGFLLLWVWADLFYFWERSRGIFLAPVFLVSPPTLLGTTLLA	120			
QY	121	TELIOLERRKGVQSSGIMTLFWLVALVICALAILRSKIMTALKEDAQVDLFRDITFYVYFS	180			
DB	121	TELIOLERRKGVQSSGIMTLFWLVALVICALAILRSKIMTALKEDAQVDLFRDITFYVYFS	180			
QY	181	LLLIQVLSCFSDRSPFLSETIHDNPPCESASFLSRITFWITGLIVRGYPQLEGSD	240			
DB	181	LLLIQVLSCFSDRSPFLSETIHDNPPCESASFLSRITFWITGLIVRGYPQLEGSD	240			
QY	241	LWSLNKEDTSEQVWPVLYVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL	300			
DB	241	LWSLNKEDTSEQVWPVLYVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL	300			
QY	301	IVKSPQKWNPSLFVLYKTFPGPYFLMSFFFKAIHDLMMFSGPQILKLIKFVNDTKAPD	360			
DB	301	IVKSPQKWNPSLFVLYKTFPGPYFLMSFFFKAIHDLMMFSGPQILKLIKFVNDTKAPD	360			
QY	361	WQGYFTVLLFVTFACLOTULVHOYFHI CFVSGMRKTAIVGAVYRKALVITNSARKSTV	420			
DB	361	WQGYFTVLLFVTFACLOTULVHOYFHI CFVSGMRKTAIVGAVYRKALVITNSARKSTV	420			
QY	421	GEIVNLSVDAORFMDLATY INNIWSAPLOVILALYLLWNLGSPVLGAVAVWMLMVPVN	480			
DB	421	GEIVNLSVDAORFMDLATY INNIWSAPLOVILALYLLWNLGSPVLGAVAVWMLMVPVN	480			
QY	481	AVYAMKTKTYQVAHMKSDNRKILMNEILNGIKLVKLYAWELAFKDKVLAIROBELKVLK	540			
DB	481	AVYAMKTKTYQVAHMKSDNRKILMNEILNGIKLVKLYAWELAFKDKVLAIROBELKVLK	540			
QY	541	KSAYLSAVGCTFTWCTPFLVALCTFAVYVTDENNILDAQTAFVSLALFNILRFPNLILP	600			
DB	541	KSAYLSAVGCTFTWCTPFLVALCTFAVYVTDENNILDAQTAFVSLALFNILRFPNLILP	600			
QY	601	MWISSIVQASVSLKRLRI FLSHEELEPDSIERPVKDGGTNSITVRNATFTWAKSDPPT	660			
DB	601	MWISSIVQASVSLKRLRI FLSHEELEPDSIERPVKDGGTNSITVRNATFTWAKSDPPT	660			
QY	661	LANGITFSIPEGALVAVVGQVCGKLSLLSALLAEMDKVEGHVAIKGSVAYVPQQAQWQND	720			
DB	661	LANGITFSIPEGALVAVVGQVCGKLSLLSALLAEMDKVEGHVAIKGSVAYVPQQAQWQND	720			
QY	721	SURENILFCQLEEPYRSVIOACALLPDLPLPSGDRTEIGEKGVLNLSGGQKQVSLAR	780			
DB	721	SURENILFCQLEEPYRSVIOACALLPDLPLPSGDRTEIGEKGVLNLSGGQKQVSLAR	780			
QY	781	AVYSNADIYLFDDPLSADAHVGHKIFENVIGPKMLKNKTRILVTHSMSYLPQVDVIV	840			
DB	781	AVYSNADIYLFDDPLSADAHVGHKIFENVIGPKMLKNKTRILVTHSMSYLPQVDVIV	840			
QY	841	MSGGKISEMGSYQELIARDGAFELRTYASTEQEQAENGVTGSGFGKEAKQWENG	900			
DB	841	MSGGKISEMGSYQELIARDGAFELRTYASTEQEQAENGVTGSGFGKEAKQWENG	900			
QY	901	LVTDSAGKQQLQSSSSSYSGDISRHNSSTAELOKAEAKBETWKLMEADQAQGVKL	960			
DB	901	LVTDSAGKQQLQSSSSSYSGDISRHNSSTAELOKAEAKBETWKLMEADQAQGVKL	960			
QY	961	SVYWDYMKAGLGFISFLSIFLFCMCHVSAASNYLSLWTDPIVNGTQHTKVRLSVYG	1020			
DB	961	SVYWDYMKAGLGFISFLSIFLFCMCHVSAASNYLSLWTDPIVNGTQHTKVRLSVYG	1020			

RESULT 14  
US-08-463-092B-6  
; Sequence 6, Application US/08463092B  
; Patent No. 5766880  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
; STREET: Queen's University at Kingston  
; CITY: Kingston  
; STATE: Ontario  
; COUNTRY: CANADA  
; ZIP: K7L 3N6  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,092B  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,893

```

; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1546
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO. 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1528 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-092B-6

Query Match 70.7%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGCADGSDPLDWNVTWNTSNPDFTKCFQNTVLVWVPCFVYLWACEPPFYFLYLSRH 60
DB 1 MALRGCADGSDPLDWNVTWNTSNPDFTKCFQNTVLVWVPCFVYLWACEPPFYFLYLSRH 60
QY 61 DRGYQMTPLNKTALGFLMIVCWADLFYSFWRSGIFLAPVFLVSPITLLGITLLA 120
DB 61 DRGYQMTPLNKTALGFLMIVCWADLFYSFWRSGIFLAPVFLVSPITLLGITLLA 120
QY 121 TPLIOLERRKGVSIGIMTFMVLVVALAILRSKIMTALKEDAOVDLFRDITFYVFS 180
DB 121 TPLIOLERRKGVSIGIMTFMVLVVALAILRSKIMTALKEDAOVDLFRDITFYVFS 180
QY 181 LLLIOLVLCFSDRPLSETIHDNCPCESSASFLRITFWITGLVIRGVRQPLEGSD 240
DB 181 LLLIOLVLCFSDRPLSETIHDNCPCESSASFLRITFWITGLVIRGVRQPLEGSD 240
QY 241 LWSLNKEDTSEGVVPLVKNWKECAKTRKQPVVYIS - SKDPAQPKSSSKVDANEVEA 299
DB 241 LWSLNKEDTSEGVVPLVKNWKECAKTRKQPVVYIS - SKDPAQPKSSSKVDANEVEA 299
QY 300 LIVSPQKEWNSLKVLYKTGPPFLMSFPFKAHDLMMFSGPOILKLLIKFVNDTKAP 359
DB 300 LIVSPQKEWNSLKVLYKTGPPFLMSFPFKAHDLMMFSGPOILKLLIKFVNDTKAP 359
QY 360 DMQGYFTYVLLFVTAQTLVHLQYFHCIFVSGMRKTAIVGAVYRKALVITNSARKSST 419
DB 360 DMQGYFTYVLLFVTAQTLVHLQYFHCIFVSGMRKTAIVGAVYRKALVITNSARKSST 419
QY 420 VGEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLWNLGSPVLGAVVWLMVVPV 479
DB 420 VGEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLWNLGSPVLGAVVWLMVVPV 479
QY 480 NAMAMKTTYQVAHMKSDNRIKLWNEITLNGIKVLKYAWELAFKDKVLAIRQBELKVL 539
DB 480 NAMAMKTTYQVAHMKSDNRIKLWNEITLNGIKVLKYAWELAFKDKVLAIRQBELKVL 539
QY 540 KKSAYLSAVGTFTWCTPFLVALCFVAVVTTDENNDLDAQAFVSLALFNILRPLNAIL 599
DB 540 KKSAYLSAVGTFTWCTPFLVALCFVAVVTTDENNDLDAQAFVSLALFNILRPLNAIL 599
QY 600 PMVISSIVQASVSLKRLIFLSHEELEPDSIERPVDGGGNSITVRNATFTWARSPP 659
DB 600 PMVISSIVQASVSLKRLIFLSHEELEPDSIERPVDGGGNSITVRNATFTWARSPP 659
QY 660 TLNGITFSPGALVAVGVGCKSSLLSALLAEMDKVEGHVAKGSAVYVPPQAWION 719
DB 660 TLNGITFSPGALVAVGVGCKSSLLSALLAEMDKVEGHVAKGSAVYVPPQAWION 719

720 DSLRENILFGQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKGVLNLSGGQKQVSLA 779
720 DSLRENILFGHPLQENYKAVMEACALLPDLLEILPSGDRTEIGEKGVLNLSGGQKQVSLA 779
780 RAVYSNADIYLFDDPLSADVAHVKGKHFENVIGPKMLKVKTRILVTHSHSYLPOVDVII 839
780 RAVYSNSDIYLFDDPLSADVAHVKGKHFENVIGPKMLKVKTRILVTHSHSYLPOVDVII 839
840 VMSGKISEMGYSQELLARDGAFAELRTYASTEQEODAEENGVTGSGPKAKOMENG 899
840 VMSGKISEMGYSQELLARDGAFAELRTYASTEQEODAEENGVTGSGPKAKOMENG 899
900 MLVTDGAKQLQRLSSSSSYSGDISRRHNSIAELQKAEAKCBETWKLMEADKAQTOQVK 959
897 MLVTDVTKHLQRLSSSSSHSGDTSQOHSSIAELQKAGA - KEETWKLMEADKAQTOQVK 955
960 LSVTYDYNKAIGLFIPISLIFLFCNHNHVSALASNYWLSLWTD - PIVNGTQEHKTVLSV 1018
956 LSVTYDYNKAIGLFIPISLIFLFCNHNHVSALASNYWLSLWTD - PIVNGTQEHKTVLSV 1015
1019 YGALGISOGLAVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFPSK 1078
1016 YGALGIIQGAARFPGYSMAVSIIGGIFASRRLHDLNVLNLRSPMSFFERTPSGNLVNRFPSK 1075
1079 ELDTVDSMIPEVIKMFMSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSR 1138
1076 ELDTVDSMIPEVIKMFMSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSR 1135
1139 QLKLESVSRSPVYSHFNETHLLGVSVIRAFEEQERFHQSDLVKDENEKAYPSIVANRW 1198
1136 QLKLESVSRSPVYSHFNETHLLGVSVIRAFEEQERFHQSDLVKDENEKAYPSIVANRW 1195
1199 LAVLECVGNCIVLFAALFAVISHRSLSAGLVLSVSLQVTTYLNLVMSSEMETNI 1258
1196 LAVLECVGNCIVLFAALFAVISHRSLSAGLVLSVSLQVTTYLNLVMSSEMETNI 1255
1259 VAVERLKEYTEKEAPWQIQETAPPSPQVGFPRNYCLRYREDLDFVLRHINVTIN 1318
1256 VAVERLKEYTEKEAPWQIQETAPPSPQVGFPRNYCLRYREDLDFVLRHINVTIE 1315
1319 GGEKVGIVGRTGAKSSITLGLFRINESABGEIIGDINTAKICLHDLRFKTIIPQDPV 1378
1316 GGEKVGIVGRTGAKSSITLGLFRINESABGEIIGDINTAKICLHDLRFKTIIPQDPV 1375
1379 LFGSLRNLDPPFQYSDSEEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGORQLV 1438
1376 LFGSLRNLDPPFQYSDSEEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGORQLV 1435
1439 CLARALLRKTKILVLDDEATAVLETDLLIOSTTRTOFEDCTVLTIAHRLNTIMDYTRVI 1498
1436 CLARALLRKTKILVLDDEATAVLETDLLIOSTTRTOFEDCTVLTIAHRLNTIMDYTRVI 1495
1499 VLDKGEIQEYCAPSDLLQORGLFYFSMAKDAGLV 1531
1496 VLDKGEVRECGAPSELQORGFIFYFSMAKDAGLV 1528

RESULT 15
US-08-462-109A-6
; Sequence 6, Application US/08462109A
; Patent No. 5882875
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; CORRESPONDENCE ADDRESS:
; NUMBER OF SEQUENCES: 6
; ADDRESS: LAHIVE & COCKFIELD
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA

```

```

;
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462.109A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966.923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029.340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141.893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407.207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.
; REGISTRATION NUMBER: 31.503
; REFERENCE/DOCKET NUMBER: PQI-002CP4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1528 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-462-109A-6

```

```

Query Match 70.7%; Score 7002.5; DB 2; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRFGCSADGSDPLWNNVTWNTSGNPDFTKCFQNTLVWVPCFYLWACFFPYFLYLSRH 60
DB 1 MALRFGCSADGSDPLWNNVTWNTSGNPDFTKCFQNTLVWVPCFYLWACFFPYFLYLSRH 60

QY 61 DRGYIQMTPLNKTALGFLMIVCWADLFYSFWERSRGIFLAPVFLVSPITLLGTTLLA 120
DB 61 DRGYIQMTPLNKTALGFLMIVCWADLFYSFWERSRGIFLAPVFLVSPITLLGTTLLA 120

QY 121 TFLQIERRKGVOSSGIMLTFWLVALCALATILRSKIMTALKEDAQVDLFRDITWVYFS 180
DB 121 TFLQIERRKGVOSSGIMLTFWLVALCALATILRSKIMTALKEDAQVDLFRDITWVYFS 180

QY 181 LLLIQLVLSFCSDRPLFSETIHDNPPCPSSASFLSRITFWIITGLIVRGYRQPLEGSD 240
DB 181 LVLVQLVLSFCSDRPLFSETIHDNPPCPSSASFLSRITFWIITGLIVRGYRQPLEGSD 240

QY 241 LWSLNKEDTSEQVVPVLVNNWKECAKTRQPKVYVS-SKDPAPQKSSSKVDANEEVEA 299
DB 241 LWSLNKEDTSEQVVPVLVNNWKECAKTRQPKVYVS-SKDPAPQKSSSKVDANEEVEA 300

QY 300 LIVSPQKWNPSLKVLYKTGPGYFLMSFFKALHDLMMFSGPQILKLIKFNVDTKAP 359
DB 301 LIVSPKHOREPSLKVLYKTGPGYFLMSFLYKALHDLMMFAGPKILELIINFVNDREAP 360

QY 360 DMQGYFTYLLVFTACLOTLVHQPHICFVSGMRIKTAVIGAVYRKALVITNSARKSST 419
DB 361 DMQGYFTYLLVFTACLOTLVHQPHICFVSGMRIKTAVIGAVYRKALVITNSARKSST 420

QY 420 VGEIVNLSVDAQRFMDLATYINNIWSAPLOVILALYLMLNLGSPSLVAGVAVMLVMPV 479
DB 421 VGEIVNLSVDAQRFMDLATYINNIWSAPLOVILALYLMLNLGSPSLVAGVAVMLVMPV 480

QY 480 NAVMAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVL 539
DB 481 NAVMAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVMSROBELKVL 540

```

Search completed: March 18, 2005, 11:03:39  
Job time : 47.2909 secs

```

QY 540 KKSAYLSAVGTFTWVCTPFLVALCTFAVYVTVIDENNILDAQTAFVSLALFNILRPPNLIL 599
DB 541 KKSAYLAAGTFTWVCTPFLVALSTFAVFTVDSERNILDAKKAFLVSLALFNILRPPNLIL 600

QY 600 PMVISSIVQASVSLKRLRIFLSHEELPDSLERPVKDGCGTNSITVRNATFTWARSPPP 659
DB 601 PMVISSIVQASVSLKRLRIFLSHEELPDSLERSIKSGEG-NSITVKNATFTWARGPP 659

QY 660 TLNGITFSIPGALVAVVGVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVVPPQAWIQN 719
DB 660 TLNGITFSIPGALVAVVGVCGCKSSLLSALLAEMDKVEGHVTLKGSVAVVPPQAWIQN 719

QY 720 DSLRENILFGQLBEPYRSVIQACALLPDLEIILPSGDRTEIGEGVNLSGQKQKRVSLA 779
DB 720 DSLRENILFGHPLQENYKAVMEACALLPDLEIILPSGDRTEIGEGVNLSGQKQKRVSLA 779

QY 780 RAVYSNADIYLFDDPLSAVDHVGKHFENYVIGPKMLKNKTRILILVTHSMYSVLPOVDVII 839
DB 780 RAVYSNSDIYLFDDPLSAVDHVGKHFEPKVPMPGLLKNKTRILILVTHGIIYSLPOVDVII 839

QY 840 VMSGKISEMGSYQELLARDGAFAEFLRTYASTEQEOQDAEENGVTGVSQPKGEAKOMENG 899
DB 840 VMSGKISEMGSYQELLARDGAFAEFLRTYASTEQEOQDAEENGVTGVSQPKGEAKOMENG 899

QY 900 MLVTSAGKQLQROLSSSSSYSGDISRHNSHTAELOKABAKKEETWKLMEADKQATGVQK 959
DB 897 MLVTDTVGKHLQRLHLSNSSSSSGDTSQOHSIABLOKAGA-KEETWKLMEADKQATGVQK 955

QY 960 LSVYNDYMKATGLFISLSEIFLFCNVHVSALASNYWLSLWTD--PUNGTQEHKTVRLSV 1018
DB 956 LSVYNYMKATGLFITFLSIFLFCNVHVSALASNYWLSLWTDPPVNVGTQANFRUSV 1015

QY 1019 YGALGISOGIAVFGYSMAVSIIGGILASRCLHVDLHLSILRSPMSPFFERTPSGNLVRFSK 1078
DB 1016 YGALGILGALAIFGYSMAVSIIGGIFASRLHLDLLYVLRSPMSPFFERTPSGNLVRFSK 1075

QY 1079 ELDTVDSMIPVIRKMFMSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSR 1138
DB 1076 ELDTVDSMIPVIRKMFMSLNFVIGAVIIILLATPIAAVIPPPLGLVYFFVQRFVASSR 1135

QY 1139 QLKLESVRSRPVYSHFNETLLGVSVTRAFSEORFTHQSDIKVDENOKAYYPSIVANRW 1198
DB 1136 QLKLESVRSRPVYSHFNETLLGVSVTRAFSEORFTHQSDIKVDENOKAYYPSIVANRW 1195

QY 1199 LAVRLECVGNCIVLFAALFAVISRHSLSAGLVGSYSLSQVTTYLNMLVRMSSEMETNI 1258
DB 1196 LAVRLECVGNCIVLFAALFAVISRHSLSAGLVGSYSLSQITAYLNMLVRMSSEMETNI 1255

QY 1259 VAVERLKEYSTEKAPMQIOETAPPSWPQVGRVPRNYCLRYREDLDFVLRHINVTIN 1318
DB 1256 VAVERLKEYSTEKAPMQIOETAPPSWPHSGRVEFRDYCLRYREDLDFVLKHINVTIE 1315

QY 1319 GGEKGVIGRTGAGKSSITLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPV 1378
DB 1316 GGEKGVIGRTGAGKSSITLGLFRINESAGEIIIDGVNIAKIGLHDLRFKTIIPQDPV 1375

QY 1379 LFSGSLRNLDPPFSQYSDDEEVMTSLAHLKDFVSALPKLDHCAEGENLSVGORQV 1438
DB 1376 LFSGSLRNLDPPFSQYSDDEEVMTSLAHLKDFVSALPKLDHCAEGENLSVGORQV 1435

QY 1439 CLARALLRKTILVLDEATAAVDLETDLIIOSTIRTOQEDCTVLTIAHRLNTIMDYTRVI 1498
DB 1436 CLARALLRKTILVLDEATAAVDLETDLIIOSTIRTOQEDCTVLTIAHRLNTIMDYTRVI 1495

QY 1499 VLDKGEIYEGAPSDLLQORGLFYSMAXDAGLV 1531
DB 1496 VLDKGEVRECCGAPSELLOQRGIFYSMAXDAGLV 1528

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 11:03:55 ; Search time 609.338 Seconds  
(without alignments)  
1045.320 Million cell updates/sec

Title: US-10-665-283-1  
Perfect score: 9903  
Sequence: 1 MALRGFCSADGSDPLMDWNV.....RSVAVAKPKFSIPDSLS 1927

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1401741 seqs, 330541175 residues  
Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	7860	79.4	1531	11	US-09-939-853A-86
2	7860	79.4	1531	16	US-10-667-891-6
3	7860	79.4	1531	17	US-10-484-577-678
4	7769	78.5	1515	11	US-09-939-853A-87
5	7037	71.1	1388	16	US-10-408-765A-1718
6	7002.5	70.7	1528	11	US-09-939-853A-88
7	4491.5	45.4	1527	11	US-09-939-853A-83
8	4491.5	45.4	1527	11	US-09-939-853A-84
9	4491.5	45.4	1527	15	US-10-295-027-1342
10	4396	44.4	1522	11	US-09-939-853A-85
11	3860.5	39.0	1548	16	US-10-667-891-1
12	3603.5	36.4	1505	16	US-10-667-891-4
13	3601.5	36.4	1499	16	US-10-667-891-3

14	3600	36.4	1542	15	US-10-363-112-4	Sequence 4, Appli
15	3600	36.4	1545	15	US-10-252-155-4	Sequence 4, Appli
16	3600	36.4	1545	15	US-10-252-155-24	Sequence 24, Appli
17	3600	36.4	1545	15	US-10-252-155-26	Sequence 26, Appli
18	3600	36.4	1545	15	US-10-252-155-28	Sequence 28, Appli
19	3600	36.4	1545	15	US-10-252-155-30	Sequence 30, Appli
20	3600	36.4	1545	15	US-10-252-155-32	Sequence 32, Appli
21	3600	36.4	1545	15	US-10-252-155-34	Sequence 34, Appli
22	3600	36.4	1545	15	US-10-252-155-38	Sequence 38, Appli
23	3600	36.4	1545	15	US-10-252-155-40	Sequence 40, Appli
24	3600	36.4	1545	15	US-10-252-155-46	Sequence 46, Appli
25	3600	36.4	1545	15	US-10-252-155-48	Sequence 48, Appli
26	3600	36.4	1545	15	US-10-363-112-2	Sequence 2, Appli
27	3600	36.4	1545	15	US-10-363-112-6	Sequence 6, Appli
28	3600	36.4	1545	15	US-10-363-112-8	Sequence 8, Appli
29	3600	36.4	1545	15	US-10-363-112-10	Sequence 10, Appli
30	3600	36.4	1545	15	US-10-363-112-12	Sequence 12, Appli
31	3600	36.4	1545	15	US-10-363-112-14	Sequence 14, Appli
32	3600	36.4	1545	15	US-10-363-112-16	Sequence 16, Appli
33	3599.5	36.3	1507	16	US-10-667-891-5	Sequence 5, Appli
34	3599	36.3	1545	15	US-10-252-155-44	Sequence 44, Appli
35	3597	36.3	1545	15	US-10-252-155-42	Sequence 42, Appli
36	3595	36.3	1545	16	US-10-741-601-369	Sequence 369, App
37	3595	36.3	1545	16	US-10-741-601-370	Sequence 370, App
38	3593	36.3	1545	15	US-10-252-155-36	Sequence 36, Appli
39	3582	36.2	1545	15	US-10-252-155-604	Sequence 604, App
40	3385.5	34.2	1503	10	US-09-792-616-3	Sequence 3, Appli
41	3385.5	34.2	1503	16	US-10-764-328-3	Sequence 3, Appli
42	3385.5	34.2	1503	17	US-10-741-600-1583	Sequence 1583, Ap
43	3289	33.2	1494	15	US-10-369-493-6697	Sequence 6697, Ap
44	3225	32.6	1573	15	US-10-369-493-6867	Sequence 6867, Ap
45	3225	32.6	1573	15	US-10-369-493-6868	Sequence 6868, Ap

ALIGNMENTS

RESULT 1  
US-09-939-853A-86  
; Sequence 86, Application US/09939853A  
; Publication No. US20040039163A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess et al.  
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-099  
; CURRENT APPLICATION NUMBER: US/09/939,853A  
; CURRENT FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: 60/228,191  
; PRIOR FILING DATE: 2000-08-25  
; PRIOR APPLICATION NUMBER: 60/267,300  
; PRIOR FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: 60/269,961  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/277,337  
; PRIOR FILING DATE: 2001-03-20  
; NUMBER OF SEQ ID NOS: 159  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 86  
; LENGTH: 1531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-939-853A-86

Query Match 79.4% ; Score 7860; DB 11; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MALRGFCSADGSDPLMDWNVTWNTSNPDTKCFQNTLVWVPCFYLWACFPFYLYLSRH 60  
Db 1 MALRGFCSADGSDPLMDWNVTWNTSNPDTKCFQNTLVWVPCFYLWACFPFYLYLSRH 60  
QY 61 DRGYQMTPLNKTATLGLFWIWCWADLFVSFWSRSRGIFLAPVFLVSPITLLGTTLLA 120

Db 61 DRGYIQMTPLNKTALGFLLLIVCWADLFYSFWRSGIFLAPVFLVSPITLLGTTLLA 120  
 Qy 121 TFLIQLERRKGVOSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Db 121 TFLIQLERRKGVOSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Qy 181 LLLIQLVLSFCSDRSPLFSETIHDNPPCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
 Db 181 LLLIQLVLSFCSDRSPLFSETIHDNPPCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
 Qy 241 LMSLNKEDTSEQVVPVVLVKNWKECAKTRKQPKVYVSSKDPKPKSSKVDANEVEAL 300  
 Db 241 LMSLNKEDTSEQVVPVVLVKNWKECAKTRKQPKVYVSSKDPKPKSSKVDANEVEAL 300  
 Qy 301 IVKSPQKWNPSLKFVLYKTFPGYFLMSPFFKAJHDLMMFSGPQILKLIKEVNDTKAPD 360  
 Db 301 IVKSPQKWNPSLKFVLYKTFPGYFLMSPFFKAJHDLMMFSGPQILKLIKEVNDTKAPD 360  
 Qy 361 WQGYFYTVLLFVTAQLTLVLHQYFHCIVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTAQLTLVLHQYFHCIVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420  
 Qy 421 GEIVNLMMSVDAQRFMDLATTYINMWSAPLOVILALYLLWNLGSPVLGAVVWMLMVPV 480  
 Db 421 GEIVNLMMSVDAQRFMDLATTYINMWSAPLOVILALYLLWNLGSPVLGAVVWMLMVPV 480  
 Qy 481 AVWAMKTKTYQVAHMKSKDNRIKLMNEILNGIKLVAYELAPKDKVLAIROBELKVLK 540  
 Db 481 AVWAMKTKTYQVAHMKSKDNRIKLMNEILNGIKLVAYELAPKDKVLAIROBELKVLK 540  
 Qy 541 KSAYLSAVGTFTWCTPPLVALCTFAVYVTDENNILDAQTAFAVSLALFNILRFPNLILP 600  
 Db 541 KSAYLSAVGTFTWCTPPLVALCTFAVYVTDENNILDAQTAFAVSLALFNILRFPNLILP 600  
 Qy 601 MVISSIVQASVLSKLRFLSHEELEPDSIERRPVKGGGTNSIIVRNATFTWARSDDPT 660  
 Db 601 MVISSIVQASVLSKLRFLSHEELEPDSIERRPVKGGGTNSIIVRNATFTWARSDDPT 660  
 Qy 661 LINGITFSIPEGALVAVGVQGGCKSSLALLAEMDKVEGHVAIKGSVAVYVPOQAWIOND 720  
 Db 661 LINGITFSIPEGALVAVGVQGGCKSSLALLAEMDKVEGHVAIKGSVAVYVPOQAWIOND 720  
 Qy 721 SIRENIFGCOLLEPYRSVIOACALLPDLILPSGDRTEIGEKNVLSGGQKQVSLAR 780  
 Db 721 SIRENIFGCOLLEPYRSVIOACALLPDLILPSGDRTEIGEKNVLSGGQKQVSLAR 780  
 Qy 781 AVYSNADILYFDPLSADVAHVGHKIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 Db 781 AVYSNADILYFDPLSADVAHVGHKIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 Qy 841 MSGGKISEMGSYQELLARDGAFELRTYVASTEQDAEENGVTGSGPGKEAKOMENG 900  
 Db 841 MSGGKISEMGSYQELLARDGAFELRTYVASTEQDAEENGVTGSGPGKEAKOMENG 900  
 Qy 901 LVTSAGKOLQRLSSSSSYSGDISRHNSSTAELQAKAEKKEETWKLMEADKAQGVKL 960  
 Db 901 LVTSAGKOLQRLSSSSSYSGDISRHNSSTAELQAKAEKKEETWKLMEADKAQGVKL 960  
 Qy 961 SVYDYMKAIGLFTSFLIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTVLSYVG 1020  
 Db 961 SVYDYMKAIGLFTSFLIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTVLSYVG 1020  
 Qy 1021 ALGISQGIAGVGYMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKEL 1080  
 Db 1021 ALGISQGIAGVGYMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKEL 1080  
 Qy 1081 DTVDNMIPEVIMFMGSLFNIVIGACIVILLATPAAIIIPPLGLIYFPVQFYVASSRQL 1140  
 Db 1081 DTVDNMIPEVIMFMGSLFNIVIGACIVILLATPAAIIIPPLGLIYFPVQFYVASSRQL 1140  
 Qy 1141 KRLESVRSRSPVYSHENETLLGVSVIRAFEEQERFTHQSDLVKDNOKAYVPSIVANRWLA 1200  
 Db 1141 KRLESVRSRSPVYSHENETLLGVSVIRAFEEQERFTHQSDLVKDNOKAYVPSIVANRWLA 1200

RESULT 2

US-10-667-891-6  
 ; Sequence 6, Application US/10667891  
 ; Publication No. US20040171024A1  
 ; GENERAL INFORMATION:

; APPLICANT: ROTH, CHARLES W.  
 ; APPLICANT: BREY, PAUL T.  
 ; APPLICANT: HOLM, INGE  
 ; APPLICANT: GRAILLES, MARINE  
 ; APPLICANT: RZHEITSKY, ANDREY  
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND  
 ; TITLE OF INVENTION: ANOPHELES  
 ; FILE REFERENCE: 03495.0294-00000  
 ; CURRENT APPLICATION NUMBER: US/10/667,891  
 ; CURRENT FILING DATE: 2003-09-23  
 ; PRIOR APPLICATION NUMBER: 60/413,469  
 ; PRIOR FILING DATE: 2002-09-26  
 ; NUMBER OF SEQ ID NOS: 76  
 ; SOFTWARE: Patent in Ver. 3.2  
 ; SEQ ID NO 6

; LENGTH: 1531  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-667-891-6

Query Match 79.4%; Score 7860; DB 16; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLWDNVTWNTSNDFTKCFONTVLVWVPCFYLWACPPFFLYLSRH 60  
 Db 1 MALRGFCSADGSDPLWDNVTWNTSNDFTKCFONTVLVWVPCFYLWACPPFFLYLSRH 60  
 Qy 61 DRGYIQTMTPLNKTALGFLLLIVCWADLFYSFWRSGIFLAPVFLVSPITLLGTTLLA 120  
 Db 61 DRGYIQTMTPLNKTALGFLLLIVCWADLFYSFWRSGIFLAPVFLVSPITLLGTTLLA 120  
 Qy 121 TFLIQLERRKGVOSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Db 121 TFLIQLERRKGVOSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Qy 181 LLLIQLVLSFCSDRSPLFSETIHDNPPCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
 Db 181 LLLIQLVLSFCSDRSPLFSETIHDNPPCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
 Qy 241 LMSLNKEDTSEQVVPVVLVKNWKECAKTRKQPKVYVSSKDPKPKSSKVDANEVEAL 300



Db 241 LWSLNKEDTSEQVFLVKNWKKCAKTRKQPKVYVSSKDPAPQPKSSKVDANEVEAL 300  
 Qy 301 IVKSPQKWNPSLFLVLYKTPGYPFLMGFFPKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKWNPSLFLVLYKTPGYPFLMGFFPKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360  
 Qy 361 WQGYFTVLLFVTTACQLTLVHLQYFHI CFVSGMRKTAIVGAVYKALVITNSARKSSTV 420  
 Db 361 WQGYFTVLLFVTTACQLTLVHLQYFHI CFVSGMRKTAIVGAVYKALVITNSARKSSTV 420  
 Qy 421 GEIVNLSMVDQRFMDLATYINMISAPLOVILALYLWNLGSPVLGAVAVMLMVPVN 480  
 Db 421 GEIVNLSMVDQRFMDLATYINMISAPLOVILALYLWNLGSPVLGAVAVMLMVPVN 480  
 Qy 481 AVAMKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKVLWELAFKDKVLAIROBELKVLK 540  
 Db 481 AVAMKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKVLWELAFKDKVLAIROBELKVLK 540  
 Qy 541 KSAYLSAVGTFTWVCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRPPNLILP 600  
 Db 541 KSAYLSAVGTFTWVCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRPPNLILP 600  
 Qy 601 MVISSIVQASVLSKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 Db 601 MVISSIVQASVLSKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 Qy 661 LNGITFTSPEGALVAVGVGCGKSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIQND 720  
 Db 661 LNGITFTSPEGALVAVGVGCGKSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIQND 720  
 Qy 721 SLRENILFGCQLEEPYRSVIOACALLPDLEILPSGDRTEIGEGKGNLSGGQKQVSLAR 780  
 Db 721 SLRENILFGCQLEEPYRSVIOACALLPDLEILPSGDRTEIGEGKGNLSGGQKQVSLAR 780  
 Qy 781 AVYSNADYILFDPLSADVAHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVLIIV 840  
 Db 781 AVYSNADYILFDPLSADVAHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVLIIV 840  
 Qy 841 MSGGKISEMGSYQELLARDGAFELRTVYASTEQODAEENGVTGSGPGKEAKOMENG 900  
 Db 841 MSGGKISEMGSYQELLARDGAFELRTVYASTEQODAEENGVTGSGPGKEAKOMENG 900  
 Qy 901 LVTDSAGKQLOQLSSSSSYSGDISRHHNSTAELQKAEAKKEETWKLMEADKAQGVKL 960  
 Db 901 LVTDSAGKQLOQLSSSSSYSGDISRHHNSTAELQKAEAKKEETWKLMEADKAQGVKL 960  
 Qy 961 SYWDYMKAIGLFISFLFPMCHVSNALSNYLSLWTDPIVNGTQEHKTVRLSYG 1020  
 Db 961 SYWDYMKAIGLFISFLFPMCHVSNALSNYLSLWTDPIVNGTQEHKTVRLSYG 1020  
 Qy 1021 ALGISQGIATVFGYSMAVSTIGGLASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
 Db 1021 ALGISQGIATVFGYSMAVSTIGGLASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
 Qy 1081 DTVDSMPEVIRKFMGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140  
 Db 1081 DTVDSMPEVIRKFMGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140  
 Qy 1141 KRLSVSRSPVYSHNETLILGVSIVRAPEEQRFHQSDLKVDENQKAYYPSIVANRWLA 1200  
 Db 1141 KRLSVSRSPVYSHNETLILGVSIVRAPEEQRFHQSDLKVDENQKAYYPSIVANRWLA 1200  
 Qy 1201 VLECVGNCIVLFAALFAVISHRSISAGLVGSVSLQVTTYLWNLVRMSEMETNIVA 1260  
 Db 1201 VLECVGNCIVLFAALFAVISHRSISAGLVGSVSLQVTTYLWNLVRMSEMETNIVA 1260  
 Qy 1261 VERLKEYSETEKAPWQIETAPPSSWPQGVREFRNYCLRYREDLDFVLRHINTVINGG 1320  
 Db 1261 VERLKEYSETEKAPWQIETAPPSSWPQGVREFRNYCLRYREDLDFVLRHINTVINGG 1320  
 Qy 1321 EXKGVIGRTGAGKSLTGLFRINSAEGEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380  
 Db 1321 EXKGVIGRTGAGKSLTGLFRINSAEGEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380

Qy 1381 SGLRMNLDPFSQYSDBEVMTSLELAHLKDFVSALPDKLDHECAGGENLSVGORQLVCL 1440  
 Db 1381 SGLRMNLDPFSQYSDBEVMTSLELAHLKDFVSALPDKLDHECAGGENLSVGORQLVCL 1440  
 Qy 1441 ARALLRKTILVLDEATAAVALDETTDDLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTILVLDEATAAVALDETTDDLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Qy 1501 DKGETQYGAQPSDILLQORGLFYSMAKDAGLV 1531  
 Db 1501 DKGETQYGAQPSDILLQORGLFYSMAKDAGLV 1531

RESULT 3

US-10-484-577-678  
 ; Sequence 678, Application US/10484577  
 ; Publication No. US20050032724A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: EPIDAUROS Biotechnologie Aktiengesellschaft  
 ; TITLE OF INVENTION: Means and methods for improved treatment of cancer based on UGRI:  
 ; FILE REFERENCE: P2285PCT-1  
 ; CURRENT APPLICATION NUMBER: US/10/484,577  
 ; CURRENT FILING DATE: 2004-01-22  
 ; PRIOR APPLICATION NUMBER: PCT/EP 02/08220  
 ; PRIOR FILING DATE: 2002-07-23  
 ; PRIOR APPLICATION NUMBER: EP 01 11 7608.8  
 ; PRIOR FILING DATE: 2001-07-23  
 ; PRIOR APPLICATION NUMBER: EP 02011710.7  
 ; PRIOR FILING DATE: 2002-05-24  
 ; NUMBER OF SEQ ID NOS: 683  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 678  
 ; LENGTH: 1531  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-484-577-678

Query Match 79.4%; Score 7860; DB 17; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSDGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPFYLILSRH 60  
 Db 1 MALRGFCSDGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPFYLILSRH 60  
 Qy 61 DRGIQMTPLNKTALGFLWIVCWADLFSFWERSRGIPLAPVFLVPTLLGTTLLA 120  
 Db 61 DRGIQMTPLNKTALGFLWIVCWADLFSFWERSRGIPLAPVFLVPTLLGTTLLA 120  
 Qy 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Db 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Qy 181 LLLIQLVLSCFSDRSPLETHDPNCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
 Db 181 LLLIQLVLSCFSDRSPLETHDPNCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
 Qy 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKQPKVYVSSKDPAPQPKSSKVDANEVEAL 300  
 Db 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKQPKVYVSSKDPAPQPKSSKVDANEVEAL 300  
 Qy 301 IVKSPQKWNPSLFLVLYKTPGYPFLMGFFPKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKWNPSLFLVLYKTPGYPFLMGFFPKAIHDLMMFSGPOLKLLIKFVNDTKAPD 360  
 Qy 361 WQGYFTVLLFVTTACQLTLVHLQYFHI CFVSGMRKTAIVGAVYKALVITNSARKSSTV 420  
 Db 361 WQGYFTVLLFVTTACQLTLVHLQYFHI CFVSGMRKTAIVGAVYKALVITNSARKSSTV 420  
 Qy 421 GEIVNLSMVDQRFMDLATYINMISAPLOVILALYLWNLGSPVLGAVAVMLMVPVN 480  
 Db 421 GEIVNLSMVDQRFMDLATYINMISAPLOVILALYLWNLGSPVLGAVAVMLMVPVN 480

481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIRQEBELKVLK 540  
481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIRQEBELKVLK 540  
541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDQAQAFVSLALFNILRFFLNILP 600  
541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDQAQAFVSLALFNILRFFLNILP 600  
601 MWISSIVQASVLSKRLRIFLSHEELEPDSIERPPVKDGGTNSITVRNATFTWASDRPT 660  
601 MWISSIVQASVLSKRLRIFLSHEELEPDSIERPPVKDGGTNSITVRNATFTWASDRPT 660  
661 LNGITFISPEGALVAVVGQVGGKSLLSALLAEWDKVEGHVAIKGSVAYVPOQAWIQND 720  
661 LNGITFISPEGALVAVVGQVGGKSLLSALLAEWDKVEGHVAIKGSVAYVPOQAWIQND 720  
721 SURENIFGQLEEPYRSVIOACALLPDLTILPSGDRTEIGEGKGNLSGGQKQVSLAR 780  
721 SURENIFGQLEEPYRSVIOACALLPDLTILPSGDRTEIGEGKGNLSGGQKQVSLAR 780  
781 AVYSNADIYLPDDPLSADVAHVGHKIFENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIV 840  
781 AVYSNADIYLPDDPLSADVAHVGHKIFENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIV 840  
841 MSGGKISEMGSYQELLARDGAFELRTVASTEQDAAEENGVTGSGGKEAKOMENG 900  
841 MSGGKISEMGSYQELLARDGAFELRTVASTEQDAAEENGVTGSGGKEAKOMENG 900  
901 LVTDAGKOLQOLSSSSSYSGDISRHNSHSTAELQAEAKKEETWKLMEADKAQTQGVKL 960  
901 LVTDAGKOLQOLSSSSSYSGDISRHNSHSTAELQAEAKKEETWKLMEADKAQTQGVKL 960  
961 SYVDYMKAIGLFISFLIFPMCHVNSALASNYWLSLWTDPIVNGTQEHKTVRLSYG 1020  
961 SYVDYMKAIGLFISFLIFPMCHVNSALASNYWLSLWTDPIVNGTQEHKTVRLSYG 1020  
1021 ALGISQGIAGVGSMAVSIIGTILASRCLHVDLLHSILSPMSFFERTPSGNLVRFSKEL 1080  
1021 ALGISQGIAGVGSMAVSIIGTILASRCLHVDLLHSILSPMSFFERTPSGNLVRFSKEL 1080  
1081 DTVDMSIPEVIMKMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFPVQRFYVASSRQL 1140  
1081 DTVDMSIPEVIMKMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFPVQRFYVASSRQL 1140  
1141 KELESVSRSPVYSHENETLLGVSVTRAEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200  
1141 KELESVSRSPVYSHENETLLGVSVTRAEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200  
1201 VRLECVGNICVILFAALFAVIRSRHSLSAGLVGLSVSYLQVTTYLNWLVRMSSEMETNIVA 1260  
1201 VRLECVGNICVILFAALFAVIRSRHSLSAGLVGLSVSYLQVTTYLNWLVRMSSEMETNIVA 1260  
1261 VERLKEYSETEKAPWQIQTAPPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320  
1261 VERLKEYSETEKAPWQIQTAPPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320  
1321 EKVGVIGRTGAGKSLTIGLFRINESAEGEIIIDGINIAKIGLHDLRKFIIIIIPQDPVLF 1380  
1321 EKVGVIGRTGAGKSLTIGLFRINESAEGEIIIDGINIAKIGLHDLRKFIIIIIPQDPVLF 1380  
1381 SGSLRMNLDPPFSQYSDDEEVTLSLELAHLKDFVSALPDKLDHECAGGNGLSVGRQQLVCL 1440  
1381 SGSLRMNLDPPFSQYSDDEEVTLSLELAHLKDFVSALPDKLDHECAGGNGLSVGRQQLVCL 1440  
1441 ARALLRKTILVLDEATAVLETDLLIQSTIRTOFEDCTVLTAHRLNTIMDYTRVIL 1500  
1441 ARALLRKTILVLDEATAVLETDLLIQSTIRTOFEDCTVLTAHRLNTIMDYTRVIL 1500  
1501 DKGEIQEYCAPSDLLQQRGLFYSMAKODAGLV 1531  
1501 DKGEIQEYCAPSDLLQQRGLFYSMAKODAGLV 1531

RESULT 4  
US-09-939-853A-87  
; Sequence 87, Application US/09939853A  
; Publication No. US20040039163A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess et al.  
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-099  
; CURRENT APPLICATION NUMBER: US/09/939,853A  
; CURRENT FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: 60/228,191  
; PRIOR FILING DATE: 2000-08-25  
; PRIOR APPLICATION NUMBER: 60/267,300  
; PRIOR FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: 60/269,961  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/277,337  
; PRIOR FILING DATE: 2001-03-20  
; NUMBER OF SEQ ID NOS: 159  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 87  
; LENGTH: 1515  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-939-853A-87

Query Match 78.5%; Score 7769; DB 11; Length 1515;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 DWNVTWNTSNDPDKFCQNTVILVWVPCFYLWACPPFYLXSRDRGVIOQTPLNKTKTA 76  
Db 1 DWNVTWNTSNDPDKFCQNTVILVWVPCFYLWACPPFYLXSRDRGVIOQTPLNKTKTA 60  
Qy 77 LGFLLLWVADVADLFYSFWERSRGIFLAPVFLVSTLIGITTLATLFIQLERRRGVSSG 136  
Db 61 LGFLLLWVADVADLFYSFWERSRGIFLAPVFLVSTLIGITTLATLFIQLERRRGVSSG 120  
Qy 137 IMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYFSLLIQLVLSCFSDRSP 196  
Db 121 IMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYFSLLIQLVLSCFSDRSP 180  
Qy 197 LFSITIDHNPCESSASFLSRITFWITGLIVRGYQPLEGSLWSINKEDTSEOVVVP 256  
Db 181 LFSITIDHNPCESSASFLSRITFWITGLIVRGYQPLEGSLWSINKEDTSEOVVVP 240  
Qy 257 LVKNWKECAKTRKQPVKVYVSSKDPQPKESSKVDANEVEALIVKSPQKWNPSLFKV 316  
Db 241 LVKNWKECAKTRKQPVKVYVSSKDPQPKESSKVDANEVEALIVKSPQKWNPSLFKV 300  
Qy 317 LYKTFGYPFLMSPFFKAIHDMPSGPQILKLLIKFVNDTKAPDWQGYFYTVLLFVTAQL 376  
Db 301 LYKTFGYPFLMSPFFKAIHDMPSGPQILKLLIKFVNDTKAPDWQGYFYTVLLFVTAQL 360  
Qy 377 QTLVLHGYFHICFVSGNRITKAVIGAVYKALVITNSARKSSTVGEIVNLSVDAQRPMD 436  
Db 361 QTLVLHGYFHICFVSGNRITKAVIGAVYKALVITNSARKSSTVGEIVNLSVDAQRPMD 420  
Qy 437 LATYINNIWSAPLQVILALYLLMLNLGSPVLAVVAVMLVMPVNAVMAKTKTYQVAHMK 496  
Db 421 LATYINNIWSAPLQVILALYLLMLNLGSPVLAVVAVMLVMPVNAVMAKTKTYQVAHMK 480  
Qy 497 SKNRIKLMNEILNGIKVLKYAWELAFKDKVLAIRQEBELKVLKKSAYLSAVGTFTWCT 556  
Db 481 SKNRIKLMNEILNGIKVLKYAWELAFKDKVLAIRQEBELKVLKKSAYLSAVGTFTWCT 540  
Qy 557 PFLVALCTFAVYVTTIDENNILDQAQAFVSLALFNILRFFLNILPMWISSIVQASVLSKRL 616  
Db 541 PFLVALCTFAVYVTTIDENNILDQAQAFVSLALFNILRFFLNILPMWISSIVQASVLSKRL 600  
Qy 617 RIFLSHEELEPDSIERPPVKDGGTNSITVRNATFTWASDRPTLNGITFISPEGALVAV 676  
Db 617 RIFLSHEELEPDSIERPPVKDGGTNSITVRNATFTWASDRPTLNGITFISPEGALVAV 676

Db 601 RIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPTTLNGITFTSIPEGALVAV 660  
 QY 677 VQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIONDSIRENIFPGCOLEBPY 736  
 Db 661 VQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIONDSIRENIFPGCOLEBPY 720  
 QY 737 YRSVTOACALLPDLBELPSGDRTEIGEKGVNLSGGQKQVSLARAVYSNADIIYLFDDPLS 796  
 Db 721 YRSVTOACALLPDLBELPSGDRTEIGEKGVNLSGGQKQVSLARAVYSNADIIYLFDDPLS 780  
 QY 797 AYDAHVGHKHFENVIGPKMKNKTRILVTHSMYSYLPQVDVIIVMSSGGKISEMGSYQELL 856  
 Db 781 AYDAHVGHKHFENVIGPKMKNKTRILVTHSMYSYLPQVDVIIVMSSGGKISEMGSYQELL 840  
 QY 857 ARDGAFAEFLRYASTEQDQAEENGVTGVSFGKEAKQKMGMLVTDTSAGKOLQRLSS 916  
 Db 841 ARDGAFAEFLRYASTEQDQAEENGVTGVSFGKEAKQKMGMLVTDTSAGKOLQRLSS 900  
 QY 917 SSSYSGDISRHNSHNTAELOKABAKKEETWKLMEADKAOTGVKLSVYWDYMKAIGLFTSF 976  
 Db 901 SSSYSGDISRHNSHNTAELOKABAKKEETWKLMEADKAOTGVKLSVYWDYMKAIGLFTSF 960  
 QY 977 LSIFLFCMCHVSALASNYWLSLWTDPIVNGTOEHTKVLRSVYGALGISOGIAVFGYSNA 1036  
 Db 961 LSIFLFCMCHVSALASNYWLSLWTDPIVNGTOEHTKVLRSVYGALGISOGIAVFGYSNA 1020  
 QY 1037 VSGIGILASRLHVDLLSHILSPMSFFERTPSGNLVRNFSKELDTVDSMIPEVIMKPMG 1096  
 Db 1021 VSGIGILASRLHVDLLSHILSPMSFFERTPSGNLVRNFSKELDTVDSMIPEVIMKPMG 1080  
 QY 1097 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPVYSHFN 1156  
 Db 1081 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPVYSHFN 1140  
 QY 1157 ETLGLSVIRAFEBQERFTHQSDLVKVDENOKAYPSIVANRWLAVRLECVGNCIVLFAAL 1216  
 Db 1141 ETLGLSVIRAFEBQERFTHQSDLVKVDENOKAYPSIVANRWLAVRLECVGNCIVLFAAL 1200  
 QY 1217 FAVISRHSLSAGLVLSVSYSLQVTTYLNWLRMSSEMETNIVAVERLKEYSETKEAPW 1276  
 Db 1201 FAVISRHSLSAGLVLSVSYSLQVTTYLNWLRMSSEMETNIVAVERLKEYSETKEAPW 1260  
 QY 1277 QIOTAPPSSWPQGVREFRNYCLARYREDLDFVLRHINVTYNGEKGKVGIVGRTGAGKSSL 1336  
 Db 1261 QIOTAPPSSWPQGVREFRNYCLARYREDLDFVLRHINVTYNGEKGKVGIVGRTGAGKSSL 1320  
 QY 1337 TLGLFRINSAEGERIIIDGINIAKIGLHDLRPKITIIIPQDPVLFSGSLRMLNLDPSQYSD 1396  
 Db 1321 TLGLFRINSAEGERIIIDGINIAKIGLHDLRPKITIIIPQDPVLFSGSLRMLNLDPSQYSD 1380  
 QY 1397 BEVWTSLELAHLKDFVSALPKDLDEHCAEGGENLSVGQRLVCLARALLRKTILVLDEA 1456  
 Db 1381 BEVWTSLELAHLKDFVSALPKDLDEHCAEGGENLSVGQRLVCLARALLRKTILVLDEA 1440  
 QY 1457 TAAVDLETDDLIOSTIRTOFECTVLTIAHRLNTIMDYTRVIVLDKGEIQEYGAPESDLLQ 1516  
 Db 1441 TAAVDLETDDLIOSTIRTOFECTVLTIAHRLNTIMDYTRVIVLDKGEIQEYGAPESDLLQ 1500  
 QY 1517 QRGLFYMAKADAGLV 1531  
 Db 1501 QRGLFYMAKADAGLV 1515

RESULT 5  
 US-10-408-765A-1718  
 ; Sequence 1718, Application US/10408765A  
 ; Publication No. US20040101874A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ghosh, Soumitra S.  
 ; APPLICANT: Fahy, Eoin D.  
 ; APPLICANT: Zhang, Bing  
 ; APPLICANT: Gibson, Bradford W.  
 ; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.  
 ; APPLICANT: Warnock, Dale E.  
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
 ; FILE REFERENCE: 660088.465  
 ; CURRENT APPLICATION NUMBER: US/10/408,765A  
 ; CURRENT FILING DATE: 2003-04-04  
 ; NUMBER OF SEQ ID NOS: 3077  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1718  
 ; LENGTH: 1388  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-408-765A-1718  
 Query Match 71.1%; Score 7037; DB 16; Length 1388;  
 Best Local Similarity 97.0%; Pred. No. 0;  
 Matches 1387; Conservative 0; Mismatches 1; Indels 42; Gaps 1;  
 QY 67 MTPLNKTKTALGFLLMIVCWADLFYSFWERSRGIFLAPVFLVSPPTLLGTTLLATFLIQL 126  
 Db 1 MTPLNKTKTALGFLLMIVCWADLFYSFWERSRGIFLAPVFLVSPPTLLGTTLLATFLIQL 60  
 QY 127 ERRKGVQSSGIMLTFWLVVALYCALAILRSKIMTALKEDAQVDLPRDITFYVYFSLLLQL 186  
 Db 61 ERRKGVQSSGIMLTFWLVVALYCALAILRSKIMTALKEDAQVDLPRDITFYVYFSLLLQL 97  
 QY 187 VLSCFSDRSPLFSETIHPNCPSSASFLSRITFWMITGLIVRGYROPLEGSDLSLNMK 246  
 Db 98 -----NPCPSSASFLSRITFWMITGLIVRGYROPLEGSDLSLNMK 138  
 QY 247 EDTSQVVPVLVKNWKECAKTRKQPVKVYSSSKDPAQPKSSKYDANEVEEALIVKSPQ 306  
 Db 139 EDTSQVVPVLVKNWKECAKTRKQPVKVYSSSKDPAQPKSSKYDANEVEEALIVKSPQ 198  
 QY 307 KEWNPSEKVLKTPGPFPLMSFFPKAIDHLMFSGPQILKLLIKFVNDTKAPDMQGYFY 366  
 Db 199 KEWNPSEKVLKTPGPFPLMSFFPKAIDHLMFSGPQILKLLIKFVNDTKAPDMQGYFY 258  
 QY 367 TVLLFVTTACLOTVLHQLVHFICFVSGMRIKTAVIGAVYRKALVITNSARKSSTVGEIVNL 426  
 Db 259 TVLLFVTTACLOTVLHQLVHFICFVSGMRIKTAVIGAVYRKALVITNSARKSSTVGEIVNL 318  
 QY 427 MSVDAQRFMDLATYINMIWSAPLOVILALYLLWNLGSPVLAGVAVMLVMPVNAVMAK 486  
 Db 319 MSVDAQRFMDLATYINMIWSAPLOVILALYLLWNLGSPVLAGVAVMLVMPVNAVMAK 378  
 QY 487 TKTQVAHMKSDNRKILMNEILNGIKVLKYAWELAPKDKVLAIRQBELKVLKKSAYLS 546  
 Db 379 TKTQVAHMKSDNRKILMNEILNGIKVLKYAWELAPKDKVLAIRQBELKVLKKSAYLS 438  
 QY 547 AVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPPLNIPMWISSI 606  
 Db 439 AVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPPLNIPMWISSI 498  
 QY 607 VQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPTTLNGITF 666  
 Db 499 VQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPTTLNGITF 558  
 QY 667 SIPEGALVAVVGVQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIONDSIRENI 726  
 Db 559 SIPEGALVAVVGVQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIONDSIRENI 618  
 QY 727 LFGCOLEBPYRYSVIOACALLPDLBELPSGDRTEIGEKGVNLSGGQKQVSLARAVYSNA 786  
 Db 619 LFGCOLEBPYRYSVIOACALLPDLBELPSGDRTEIGEKGVNLSGGQKQVSLARAVYSNA 678  
 QY 787 DIYLFDDPLSADVAHVGHKHFENVIGPKMKNKTRILVTHSMYSYLPQVDVIIVMSSGGKI 846  
 Db 679 DIYLFDDPLSADVAHVGHKHFENVIGPKMKNKTRILVTHSMYSYLPQVDVIIVMSSGGKI 738  
 QY 847 SEMGSYQELLARDGAFABFLRTYASTEQDQAEENGVTGVSFGKEAKQKMGMLVTDTSAG 906

Db 739 SEMGYSQELLARDGAFELRTYASTEQEADAEENGVTGSGPGKEAKQWENGLVTD5A 798  
 Qy 907 GKOLQRLSSSSYSYSDISRHNSTAELOKAEAKKEETWKLMEADKAQOGVYKLSYWDY 966  
 Db 799 GKOLQRLSSSSYSYSDISRHNSTAELOKAEAKKEETWKLMEADKAQOGVYKLSYWDY 858  
 Qy 967 MKAIGLIFSLSTIFMCHNVHVSALASNYWLSWTDDPIVNGTQETHKVRLSVYGALGISQ 1026  
 Db 859 MKAIGLIFSLSTIFMCHNVHVSALASNYWLSWTDDPIVNGTQETHKVRLSVYGALGISQ 918  
 Qy 1027 GIAVGYSMVAGIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKELPTVDSM 1086  
 Db 919 GIAVGYSMVAGIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKELPTVDSM 978  
 Qy 1087 IPEVIMKMGSLFNIVIGACIVILLATPAAIIIPGLIYFFVQVRYVASSRQLKRLSV 1146  
 Db 979 IPEVIMKMGSLFNIVIGACIVILLATPAAIIIPGLIYFFVQVRYVASSRQLKRLSV 1038  
 Qy 1147 SRSPVSHNETLLGVSIVRAFEEOERFTHQSDLVKVDENOKAYPSIVANRWLAVRLECV 1206  
 Db 1039 SRSPVSHNETLLGVSIVRAFEEOERFTHQSDLVKVDENOKAYPSIVANRWLAVRLECV 1098  
 Qy 1207 GNCIVLFAALFAVISRHSLSAGLVGSYSYSLQVTTYLNWLVRMSSEMETNIVAVERLKE 1266  
 Db 1099 GNCIVLFAALFAVISRHSLSAGLVGSYSYSLQVTTYLNWLVRMSSEMETNIVAVERLKE 1158  
 Qy 1267 YSETEKEAPWQIOETAPSSPQVGRVFRNVCYLYREDLPVLRHINVTINGGKVGIV 1326  
 Db 1159 YSETEKEAPWQIOETAPSSPQVGRVFRNVCYLYREDLPVLRHINVTINGGKVGIV 1218  
 Qy 1327 GRTGAGKSSLTGLFRINESAGEIIDIINIAGIHLDRPKITIIIPQDPVLFGSLRM 1386  
 Db 1219 GRTGAGKSSLTGLFRINESAGEIIDIINIAGIHLDRPKITIIIPQDPVLFGSLRM 1278  
 Qy 1387 NLDPSQYSDDEWVTSLELAHLKDFVSALPKLDHECAEGGENLSVGQRLVCLARALLR 1446  
 Db 1279 NLDPSQYSDDEWVTSLELAHLKDFVSALPKLDHECAEGGENLSVGQRLVCLARALLR 1338  
 Qy 1447 KTKILVLDEATAVLETDLLIQSTIRTOFREDCTVLTIAHRLNTIMDYTR 1496  
 Db 1339 KTKILVLDEATAVLETDLLIQSTIRTOFREDCTVLTIAHRLNTIMDYTR 1388

RESULT 6  
 US-09-939-853A-88  
 ; Sequence 88, Application US/09939853A  
 ; Publication No. US20040039163A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Burgess et al.  
 ; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same  
 ; FILE REFERENCE: 21402-099  
 ; CURRENT APPLICATION NUMBER: US/09/939,853A  
 ; PRIORITY FILING DATE: 2001-08-27  
 ; PRIOR APPLICATION NUMBER: 60/228,191  
 ; PRIOR FILING DATE: 2000-08-25  
 ; PRIOR APPLICATION NUMBER: 60/267,300  
 ; PRIOR FILING DATE: 2001-02-08  
 ; PRIOR APPLICATION NUMBER: 60/269,961  
 ; PRIOR FILING DATE: 2001-02-20  
 ; PRIOR APPLICATION NUMBER: 60/277,337  
 ; PRIOR FILING DATE: 2001-03-20  
 ; NUMBER OF SEQ ID NOS: 159  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 88  
 ; LENGTH: 1528  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 US-09-939-853A-88

Query Match 70.7%; Score 7002.5; DB 11; Length 1528;  
 Best Local Similarity 88.0%; Pred. No. 0;  
 Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

Qy 1 MALRGFCADSGDPLDWNVTWNTSNPDFTKCFONTVLVWVPCFVLWACFPFFFLYLSRH 60  
 Db 1 MALRGFCADSGDPLDWNVTWNTSNPDFTKCFONTVLVWVPCFVLWACFPFFFLYLSRH 60  
 Qy 61 DRGVIQMTPLNKTATGALGFLIWCWADLFYSFWERSRGIFLAPVFLVPSYLLGTTLLA 120  
 Db 61 DRGVIQMTPLNKTATGALGFLIWCWADLFYSFWERSRGIFLAPVFLVPSYLLGTTLLA 120  
 Qy 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITPYVYFS 180  
 Db 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITPYVYFS 180  
 Qy 181 LLLIQLVLSCSFSPSETHDNPCESSASFLSRIITFWMTGLVIRGVRQPLESSD 240  
 Db 181 LLLIQLVLSCSFSPSETHDNPCESSASFLSRIITFWMTGLVIRGVRQPLESSD 240  
 Qy 241 LWSLNKEDTSQVVPVVLVKNWKECAKTRKQPVVWVS-SKDPAPQPKESSKVDANEVEEA 299  
 Db 241 LWSLNKEDTSQVVPVVLVKNWKECAKTRKQPVVWVS-SKDPAPQPKESSKVDANEVEEA 299  
 Qy 300 LIVSPQKEMNPSPFKVLYKTFGPFYFMSFFKAIHDLMPSPGOILKLLIKFVNDTKAP 359  
 Db 301 LIVSPKHDRPSPFKVLYKTFGPFYFMSFLYKALHDLMPGPKILELIINFVNDREAP 360  
 Qy 360 DWQGVYTVLLFVTACIQLTLVHOYEHICFVSGMRKIKTAVIGAVYRKALVTNSAKSST 419  
 Db 361 DWQGVYTVLLFVTACIQLTLVHOYEHICFVSGMRKIKTAVIGAVYRKALVTNSAKSST 420  
 Qy 420 VGEIVNLMSVDAORFMDLATYINNIWSAPQVILALYLLMLNLGSPVLAVVAVMLVVPV 479  
 Db 421 VGEIVNLMSVDAORFMDLATYINNIWSAPQVILALYLLMLNLGSPVLAVVAVMLVVPV 480  
 Qy 480 NAVMAMTKTYQVAHMKSKONRIKLMNEILNGIKVLKYAWELAFKQVLAIRQELKVL 539  
 Db 481 NAVMAMTKTYQVAHMKSKONRIKLMNEILNGIKVLKYAWELAFKQVLAIRQELKVL 540  
 Qy 540 KKSAYLSAVGTFVWCTPFLVALCTFAVYVTDENNILDAQTAQVSLALFNILFPNLIL 599  
 Db 541 KKSAYLSAVGTFVWCTPFLVALCTFAVYVTDENNILDAQTAQVSLALFNILFPNLIL 600  
 Qy 600 PMVISSIVQASVSLKRLRIFLSHHELEPDSIERRPVKDGGTNSITVRNATFTTWARSDPP 659  
 Db 601 PMVISSIVQASVSLKRLRIFLSHHELEPDSIERRSIKSGEG-NSITVKNATFTTWARGEPP 659  
 Qy 660 TLNGITFSIPEGALVAVVGVQCGCKSSLSALLAEMDKVEGHVAIKSVAVVPOQAIQN 719  
 Db 660 TLNGITFSIPEGALVAVVGVQCGCKSSLSALLAEMDKVEGHVTLKGSVAVVPOQAIQN 719  
 Qy 720 DSLRENILFGQLEPEYRSVIOACALLPDLLEILPSGDRTEIGKGVNLSGGQKQVSLA 779  
 Db 720 DSLRENILFGHPLQENYKAVMEACALLPDLLEILPSGDRTEIGKGVNLSGGQKQVSLA 779  
 Qy 780 RAVYSNADIYLFDDPLSAVDAHVGKHFENVIGPKGMLKNKTRILVTHSHSYLPOVDVII 839  
 Db 780 RAVYSNADIYLFDDPLSAVDAHVGKHFENVIGPKGMLKNKTRILVTHSHSYLPOVDVII 839  
 Qy 840 VMGGKISEMGYSOELLARDGAFELRTYASTEQEADAEENGVTGSGPGKEAKQWENGLV 899  
 Db 840 VMGGKISEMGYSOELLARDGAFELRTYASTEQEADAEENGVTGSGPGKEAKQWENGLV 896  
 Qy 900 MLVTDTSAGKQLQRLSSSSYSYSDISRHNSTAELOKAEAKKEETWKLMEADKAQOGVY 959  
 Db 897 MLVTDTSAGKQLQRLSSSSYSYSDISRHNSTAELOKAEAKKEETWKLMEADKAQOGVY 955  
 Qy 960 LSVYWDYMKAIGLIFISFLSIFLFWCNHVSALASNYWLSWTDD-PIVNGTQETHKVRLSV 1018  
 Db 956 LSVYWDYMKAIGLIFISFLSIFLFWCNHVSALASNYWLSWTDDPPVYVNGTQETHKVRLSV 1015  
 Qy 1019 YGALGISQIAGVAGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSK 1078  
 Db 1016 YGALGISQIAGVAGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSK 1075  
 Qy 1079 ELDTVDSMIPEVIMKMGSLFNIVIGACIVILLATPAAIIIPGLIYFFVQVRYVASSR 1138













Db 1365 QDPIFGSLTLMNLDPEGRYSEDEIWRTELHSLHSAFVSQPTGLDFQCSEGGDNLVSG 1424  
 Qy 1434 QRLVCLARALKRKILVLDATAAVALDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMD 1493  
 Db 1425 QRLVCLARALKRKILVLDATAAVALDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMD 1484  
 Qy 1494 YTRVILDKGEIQEYGAPELQORGLFLYMAKADAGL 1530  
 Db 1485 YNRVLVDKGVVAEPDPSPNLTAAGGIFPYMAKADAGL 1521

RESULT 11  
 US-10-667-891-1  
 ; Sequence 1, Application US/10667891  
 ; Publication No. US20040171024A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ROTH, CHARLES W.  
 ; APPLICANT: BREY, PAUL T.  
 ; APPLICANT: HOLM, INGE  
 ; APPLICANT: GRAILLES, MARINE  
 ; APPLICANT: RZHETSKY, ANDREY  
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND  
 ; TITLE OF INVENTION: ANOPHELES  
 ; FILE REFERENCE: 03495.0294-00000  
 ; CURRENT APPLICATION NUMBER: US/10/667,891  
 ; CURRENT FILING DATE: 2003-09-23  
 ; PRIOR APPLICATION NUMBER: 60/413,469  
 ; PRIOR FILING DATE: 2002-09-26  
 ; NUMBER OF SEQ ID NOS: 76  
 ; SOFTWARE: Patent in Ver. 3.2  
 ; SEQ ID NO 1  
 ; LENGTH: 1548  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 ; US-10-667-891-1

Query Match 39.0%; Score 3860.5; DB 16; Length 1548;  
 Best Local Similarity 50.0%; Pred. No. 2e-314;  
 Matches 784; Conservative 274; Mismatches 438; Indels 71; Gaps 18;  
 Qy 6 FCSAGSDPLDWNVTWNTSNDFTKCFQNTVLVWPCFLWA--CFPPYFLYLSHRBG 63  
 Db 12 FCGS----TFMNATETWYTNDDPTCFRQTALVWTPCAFYWAFFDFYFLKASL-DRN 66  
 Qy 64 YIQMTPLNKTALGFLWVWADLFYSEWRSRGIFLAPVF--LVSPTELLGTTLLA 120  
 Db 67 -IPWNKLVNKAIVNGLLVITALLDIMALVKKG-GDSELPYLDLDVWGPITKPAITLL 124  
 Qy 121 TFLIQLERKGVQSSGIMLTFWLVALVCAIALILRSKI-----MTALKEDAQVD 168  
 Db 125 FIFPLNRKYGVQTTGCFIFWFLTLVLSIPRCRTEVRLDAERQKILNSQQPDSFWE 184  
 Qy 169 LPRDTFYVYFLLLIQLVLSFSDRSPLFSETIHDNPNCPRESSASFLSRIITFWITGLI 228  
 Db 185 EYQVFSFFFTFTSIMLILNCFADGMPRTQYQGENEIPELSASFLSRITYQWFDKMA 244  
 Qy 229 VSGYQPLSGDLMSLNKEDTSEQVPLVKNWKECATKQPKVYVSSKDPAPKES 288  
 Db 245 LKGYRNPLEEKLDLDURPDQSCSEVNPFIHANNVNRKNYKNKAV-----BPKAQFNG 300  
 Qy 289 SKVDANEVEALIVKSPQEWNPSPFKVLYKTFPGPYFLMSFFPKAIHDLMPFSGQILKL 348  
 Db 301 NVTFFENPHGE----KNGRKKGWASIMPPIYKSGGVFLFGALMKLFTDLTFAQPVLSL 356  
 Qy 349 LIKPVNDTKA-PDWQGYFTVLLFVTACIQLVHOYFHI CFVSGMRKIKTAVIGAYRKA 407  
 Db 357 IISFVEAQDAEPKWKILYAVLLFVLAQAQTFILGQYFHRMFTVGLRIRITALINAIYRKA 416  
 Qy 408 LVITNSARKSSTVGEIVNLSVDAQRFMDLATVINMWISAPLOVIALVLLNIGPSVL 467  
 Db 417 LRISNTKKESTVGEIVNLSVDAQRFMDLATVINMWISAPLOVIALVLLNIGPSVL 476  
 Qy 468 AGVAVMVLVPVNAVMAKTKTKYQVAHMKSKONRIKLMNEILNIGIKVLKYLAWELAFKDK 527

Db 477 AGLAVMILIPVNGVIAIRIKTYQIRQMKYXDERVKLMEVLSGKVLKLVANEPSPFKQ 536  
 Qy 528 VLAIHQBELKVLKSAVLSAVGTTWCTPFLVALCTFAVYVTDENNILDAQAFVSLA 587  
 Db 537 VLDIRDEIATLRSTAYLNAGTSFLWSCAPFLVSLVTFATVYLDENNVLDAKTFFVSL 596  
 Qy 588 LFNILRFPNLPLMVISIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVR 647  
 Db 597 LFNILRFPNLPLMILTNVQTVSVNRINKFLNSELDPNV-----LHDSKPHPMSTE 652  
 Qy 648 NATTTWASDPPTLNGITFTSIPGALVAVVQVCGGKSSLLSALLAENKDVGEHVAIKGS 707  
 Db 653 NGEFSW--GDEITLRNINIEVKKSLVALVGTGSGKSSVQVAFGEKELAGVNTVCK 710  
 Qy 708 VAYVPOQAWIQNDSLRNIFPGCOLEEPYRSVIOACALLPDLLEILPDSGRTEIGEGVN 767  
 Db 711 LAYVPOQAWIQNATVRDNLFGQYDVKRYNKVIDACALRADIDILSAGDLTEIGEGIN 770  
 Qy 768 LSGGQQRVSLARAVYSNADIYLFDDPLSAVDHVGKHFENVIGPKMKLNKTRILVTH 827  
 Db 771 LSGGQQRVSLARAVYSNADIYLFDDPLSAVDHVGKHFENVIGPKMKLNKTRILVTH 830  
 Qy 828 SMSYLPQVDVLIIVMSGKISSEMSYQELLARDGAFABFLRTYASTEQEQDAEEN----- 881  
 Db 831 GVTFLPQVDSIYVIMGSEISGTFDQLVKNKGAPADFIIOHLOQSGNEBSEBELNQIKRQI 890  
 Qy 882 ---GVTGSGPGKEAKOMENGLVTD-----AGKQORQLSSSSSYSGDISR 926  
 Db 891 SSTADVPELLGTVEKAIKLARTESLSDISVTSADSLMGGGSLRRRTKQDS----- 943  
 Qy 927 HNSTAELOKAEKKEETMKLMEADKAQTKQVLSVYWDYMKAIGLFISFISIFMCMH 986  
 Db 944 -HDSVASAASLKXQEVGKLIETEKSTQGVFAVYKHYIKSVGIFLSVATLVLFVQ 1002  
 Qy 987 VSALASNYLSLWTDPIVNGTQHTKVRLSVYGALGISQIATVAVGYSMAVSGIGLASR 1046  
 Db 1003 AFQISNMLTQWANDQNVANDTGLRDMYLVGYGAGFGQVLSKYLGLALAGLHSCM 1062  
 Qy 1047 CLHVDLHLSILRSPMSFPERTPSGNLVNRFSKELDTVDMSIPEVTKMPGSLFNIGACI 1106  
 Db 1063 NVFNKLLTGKWPMEFLDTPGLRILSRYSKQVDTVDVSLPAITVQLLNTCFGLVATIV 1122  
 Qy 1107 VILLATPAAIIPPLGLIYFFVQYVASSRQKRLSVSRSPYSHFNELTLLGVSVIR 1166  
 Db 1123 VISLSTPFLAVIVPIAFLYVFAQRFYVATSRQMLRSLSVSRSPYSHFSETVTCASTIR 1182  
 Qy 1167 APEQERFIHOSDLKVDENQKAYPSIVANRWLAVRLECVGNCIVLFAALPAVISHSL 1226  
 Db 1183 AYNVGDREIIESDAKVDKNQVCKYPSVIANRWLAIRLEVMVGNLIILFASLFAVLOGQT-N 1241  
 Qy 1227 AGLVGLSVSYLSQVTTYLNWLVRMSSEMETNIVAVERLKEYSETEKEAPWQI-QETAPPS 1285  
 Db 1242 PGLVGLSVYALQVOTLNLVVRMSDDETNIIVSVVERKEYGETQKEAPWELQDKKPK 1301  
 Qy 1286 SPMQVGRVFRNYCLYREDLDFVLRHINVTINGCEKVGIVGRTGAGKSSLTGLFRINE 1345  
 Db 1302 NMPQGRVFEFQVRYREGDLVLRGVSFNIOGGEKVGIVGRTGAGKSSLTGLFRINE 1361  
 Qy 1346 SAGBIIIDGINIAKIGHDLRFPKTIIPDDPVLFGSLRMLNLDPSQVSDREVMTSEL 1405  
 Db 1362 AAGGRISIDGVDIASMGLHMLRSLTIIPDDPVLFGSLRMLNLDPSQVSDREVMTSEL 1421  
 Qy 1406 AHLKDFVSALPDKLHCEAEGENLSVGORQLVCLARALLRKTILVLDATAAVALDLETD 1465  
 Db 1422 SHLKSFKVSLAAGLHNEIAEGENLSVGORQLVCLARALLRKTILVLDATAAVALDLETD 1481  
 Qy 1466 DLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVILDKGEIQEYGAPELQORGLFLYMAKADAGL 1524  
 Db 1482 DLIQSTIRTOFEDCTVLTIAHRLNTILSDKVIDKQIIEFASPTTELLDNPKSAFYSM 1541  
 Qy 1525 AKDAGLV 1531

```

Db      1542 AKDANLV 1548

RESULT 12
US-10-667-891-4
; Sequence 4, Application US/10667891
; Publication No. US20040171024A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, CHARLES W.
; APPLICANT: BREY, PAUL T.
; APPLICANT: HOLM, INGE
; APPLICANT: GRAILLES, MARINE
; APPLICANT: RZHEITSKY, ANDREY
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
; FILE OF INVENTION: ANOPHELES
; FILE REFERENCE: 03495.0294-00000
; CURRENT APPLICATION NUMBER: US/10/667,891
; CURRENT FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: 60/413,469
; PRIOR FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 4
; LENGTH: 1505
; TYPE: PRT
; ORGANISM: Anopheles gambiae
US-10-667-891-4

Query Match      36.4%; Score 3603.5; DB 16; Length 1505;
Best Local Similarity 48.0%; Pred. No. 8e-293;
Matches 750; Conservative 280; Mismatches 444; Indels 87; Gaps 22;

QY      1 MALRFGCSADGSDPLDNDVNTWNTNPDFTKCFQNTLVWVPCFVLMACFPYFYL-LSR 59
Db      1 MTFEPFCG-----PFWDGSEFVWDVNDPNLTFCQFVILQWVCLFLV-FSIYDIFKITE 55

QY      60 HDRGYIOMTPLNKTALGFLIWIVCWADL-PYSFWERSRGIFLAPVLVSPTLGITTLL 118
Db      56 SKYRDI PNWYNLSRMLVFLMCMCWDILGMVGVQDEQGLYDVQILTAVENALAYIDL 115

QY      119 LATFLIQLERRKGVSSGIMLTFWLVALCALAILRSKIM-----TALKEDAQVDL--P 170
Db      116 LV--LLFFMRKYGVRTSGTMFWMFWRFFGIIQLRTEVMENDKRPNAIGSGDVTDFWEY 173

QY      171 RDITFVYFSLILLIOLVLSGSDRPLPSETIHDNPPCPSSASFLSRITETWITGLIVR 230
Db      174 QVSVILQYSLICMLVLELPDKPSYSPYPAKPNELRSFFSKLLFLHFDPAFAWK 233

QY      231 GYRQPLEGSLWSLNKEDTSEQVPLVKNMKCEKATRKQPVKVYVYSSKDPAPKESK 290
Db      234 GPRNPLTMNDYDINPDQSARELVPPDKYWKISVEKGRKQOM-----ASDRKAGK 284

QY      291 VDANEEVALIVKSPQKWNPSLFLVLYKTF-----GPYFLMSFFPKATHDLMFSGPQIL 346
Db      285 PDID-----YKPHSPNSGSLVLTMTIRAYGGPFWAGMLQLAISGL-QFASPYLM 332

QY      347 KLLIKFVNDTKAPDQGYFYTVLLFVTACLOTVLHGVFHCIFVSGMFIKTAIVGAVYRK 406
Db      333 QELMAVI-AFDGPLWKGFLLFGLFGLSILLGLFNGQVLYFTYFISGFRIRGLISAIYRK 391

QY      407 ALVITNSARKSTGVEIVNLMSVDAQRFMDLATYINMTWSPLOVILALYLLMLNLGPSV 466
Db      392 ALRISSAAKDDTVGEIVNLMAVDAQKFELTSLYHLILWSALLIIGLCVFLLYDILGPV 451

QY      467 LAGVAVMYLVNVAWMAKTKTQVAMHMSKDNRIKMLNELLNGIKVLKLYANELAFKD 526
Db      452 FAGLGVMLTMTVPVSGVAAKLUKTHQVAMKLUDEKVKONKILGGIKVLKLYAWEPSPQD 511

QY      527 KVLAIROBELKVLKSKAYSLSAVGTFTVCTPPLVALCTFAVYVITDENNILDQAOTAFVSL 586
Db      512 SILNVRDEEVGLKQWYAGIIFTFTIAFLVTLVSVFAYVILDENNVLDPDQTAFLVSL 571

QY      597 ALFNILRPLNPLMNVISSIVQASVSLKRLRIFLSHEELEPDSTERRFPVKDGGGTNSITY 646

```

RESULT 13  
 US-10-667-891-3  
 ; Sequence 3, Application US/10667891  
 ; Publication No. US20040171024A1  
 ; GENERAL INFORMATION:

; APPLICANT: ROTH, CHARLES W.  
 ; APPLICANT: BREY, PAUL T.  
 ; APPLICANT: HOLM, INGE  
 ; APPLICANT: GRAILLES, MARINE  
 ; APPLICANT: EZHETSKY, ANDREY  
 ; TITLE OF INVENTION: MUTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND  
 ; TITLE OF INVENTION: ANOPHELES  
 ; FILE REFERENCE: 03495.0294-00000  
 ; CURRENT APPLICATION NUMBER: US/10/667,891  
 ; CURRENT FILING DATE: 2003-09-23  
 ; PRIOR APPLICATION NUMBER: 60/413,469  
 ; PRIOR FILING DATE: 2002-09-26  
 ; NUMBER OF SEQ ID NOS: 76  
 ; SOFTWARE: Patent In Ver. 3.2  
 ; SEQ ID NO 3  
 ; LENGTH: 1499  
 ; TYPE: PRT  
 ; ORGANISM: Anopheles gambiae  
 ;  
 ; US-10-667-891-3  
 ;  
 Query Match 36.4%; Score 3601.5; DB 16; Length 1499;  
 Best Local Similarity 47.9%; Pred. No. 1-2e-292;  
 Matches 744; Conservative 283; Mismatches 448; Indels 79; Gaps 20;  
 ;  
 QY 1 MALRGCSADGSDPLMDWNTVNTSNPDTKCFQNTLVVWPCFVLMACFPFYLYLSRH 60  
 DB 1 MTFEEFCGG-PPWDDDLTWREEDPDLTFCFQVILQWTFQFLFVFSMVEVLRIVTS 56  
 QY 61 DRGYQMTPLNKTALGALLMIVCWADLFYFWSERSGIFLAPVLPVSPITLIGIT- 116  
 DB 57 RYRDIPWNNFNITKMTFTALMVMSWDL-----GVGLANDWIES--LLSMLPSLP 105  
 QY 117 -TLATFLQLERKQVQSGQWMLTFWLVVALCALAILRSKWT-ALKEDAQVLD--PRD 172  
 DB 106 FOIMAMALVFFRYKYGIRGTGTFIFWFLKAFGIQMTKTEAMLDHVRSGTGDAEFOF 165  
 QY 173 ITFYVYFSLLLIQLVLSCFSDRSPLESTHDPNCPSSAGFLSKITFWITGLIVRGY 232  
 DB 166 VSYTIQYVCCVLLLELFPDKREPRYSEWAKLNPNPRLSSFFSRLFYLYFDSVAMRGF 225  
 QY 233 RQPLEGSDLSLNKETSQVVPVLVWKKCEKATKRPVVKVYSSKDPQPKSSKVD 292  
 DB 226 RKPLTDDMDYLNPEDTSAALVPPDPKYYESVEKGRKQI-----AADKAGKTN 276  
 QY 293 ANEEVEALVKSPOKEWNSLPKVLKTP-GPYFLMSFFFKATHDLMFSGPOILKLLIK 351  
 DB 277 -----LVYKNAATNGSVLPAMVKA YGCFWFAQMLQPAISGL-QFASPYLQEIWA 327  
 QY 352 FVNDTKAPWQGYFTVLLFTACQLTLVLHQYFHCYFVGMRITKAVIGAVYRKALVIT 411  
 DB 328 VI-ALDGPFWKGMIIITGLFLSLIALFNGQYFHTFLVGFRIQTGLISAIYRKALRIS 386  
 QY 412 NGARKSTGEIVNLSVDAQRPMDLATYINMWSAPLOVILALYLLWNLGSPVLAVGA 471  
 DB 387 SFAKDQTVGEIVNLSVDAQRPMDLATYINMWSAPLOVILALYLLWNLGSPVLAVGA 446  
 QY 472 VVWLVVPAVNAWMAKTKTVQAHMKSNDNRILKMLNEILANGIKVLKYAWELAFKDKVLAI 531  
 DB 447 VVWIMIPITGFIATRDIDQVEQMKIKDERVKQKNEILGKIKVLYAWEPFQDITVTV 506  
 QY 532 ROBELKVLKKSAYLSAVGFTTWYCTPFLVALCTFAVYVTTIDENNLIDQAQTAFLVSLAFNI 591  
 DB 507 RNEELVLSAAVYAGTGFVWMTWAPFLVTLASFAVYVMDIENVLDPQTAFLVALFNI 566  
 QY 592 LRPLNIPMVISSIVQASVSLKRLRIFLSHBELEPDSTERRPKDGGGTNSITVFNATP 651  
 DB 567 LRPLAMFPMITFAMQAWYSIKRIDKFMNSELDNNVTHNK-----SENALVXKDTGF 621  
 QY 652 TWARSDPPTLNGITFSIPGALVAVVGVCGCKSSLLSALLAEMDKVGHVAKGSAVY 711  
 DB 622 SWG-DDAPTLKTNLNRGKUSAVVGVGTGKSSLSIALGMEKMGKSVNTDGSIAV 680  
 QY 712 PQQAWIQNSLRNLFQCGLEBPYYSRVIQACALLPDLEILPSPGDRTEIGEGVNLGSG 771

DB 681 PQQAWIQNSLRNLFQCGLEBPYYSRVIQACALLPDLEILPSPGDRTEIGEGVNLGSG 740  
 QY 772 OKQVSLARAVVSNADIVLFDPLSAVDAHVGKHFENVIGPKMLKKNKTRILVTHSMYSY 831  
 DB 741 QKQVALLARAVYADSEVYLFDDPLSAVDAHVGKHFENVIGPKMLKKNKTRILVTHSIF 800  
 QY 832 LPQDVIIIVMGSGKISEMGSYOELLARDAFABFRTYASTEQEQAENGVTGSGPK 891  
 DB 801 LPFVEEIFVMKDGESGSGYOELLARDAFABFRTYASTEQEQAENGVTGSGPK 848  
 QY 892 EAKQWENGMLVYDSAGKQIQRQLSSSSSYSGDI-----SRHNSAELOKAE 938  
 DB 849 ELKIQEA-LKDGEAKTVQRAMSTRSQSSNGSVRKRVSRAESRNSKPRAVEQTV 907  
 QY 939 AKKEETWKLMEADKAQTOGVKLSVYWDYMKATGLFISFLSIFLFCNHYVSALASNYLSL 998  
 DB 908 AQSSA-TLIEKESATGAVGVVYIKYFGKGLWGLFWISFFSVINQOASIVANIWLTD 966  
 QY 999 WTDDPIVNGTQHTKVRLSVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1058  
 DB 967 WSEDEPAATDPSVRDWMYLVGGLGGAQSIALLIASVTALALGCIKAARELHNNLESSNR 1026  
 QY 1059 SPMSFERTPSGNLVNRSKELDTVDSMIPEVVKFMGSLFNVIACIVILLATPAAII 1118  
 DB 1027 MPMSPFDTPLGRIMNRFKSDVVDNILPQSIKAWLLMFFNVGVVYVYVYVYVYVYVY 1086  
 QY 1119 IPLGLIYFFQVY 1178  
 DB 1087 VPFLVYIY 1146  
 QY 1179 DLKVDENKAYPSYVIANRWLAVRLECVGNCIVLFAALPAVISRHSLSAGLVSVSYSL 1238  
 DB 1147 EQVDYNQLTYSYPIANRWLAVRLECVGNCIVLFAALPAVISRHSLSAGLVSVSYSL 1206  
 QY 1239 QVTTVNLWLMVMSSEMETNIVAVRLKEYSETEKAPQIQETAPPSWPOVGRVFRNY 1298  
 DB 1207 QISATSLFLVRMTAEVETNIVAIERLEBYTVLPREAWQLGHV--DKAMPVEKGVFKDY 1264  
 QY 1299 CLURYREDLDFVLRHINVTINGEKVIGVORTGAGKSSLTGLGLFRINESAGEIIIDGINI 1358  
 DB 1265 QIRYREGDLVIRGSLNVRGKEKIGVORTGAGKSSLTGLGLFRINESAGEIIIDGINI 1324  
 QY 1359 AKIGHDLRFKTIITIPQDPVLPFSGSLRWMLDPFSGYSDVEVWTSLELAHLKDFVSALPK 1418  
 DB 1325 SKWGLHQLRGLTIIIPQDPVLPFSGSLRWMLDPFSGYSDVEVWTSLELAHLKDFVSALPK 1384  
 QY 1419 LDHECAEGENLSVQORQLVCLARALLKTKILVILDEATAAVDLETDLLIQSTIRTFQED 1478  
 DB 1385 LDHECAEGENLSVQORQLVCLARALLKTKILVILDEATAAVDLETDLLIQSTIRTFQED 1444  
 QY 1479 CTVLTIAHRLNTIMDYTVRIVLDKGEIOEYGAAPSLLQOR-GLFYSMKADAGLV 1531  
 DB 1445 CTILTIAHRLNTILSDRVLDKGLVAECDSFQNLNANRESIFPGMAKNAGIV 1498

RESULT 14  
 US-10-363-112-4  
 ; Sequence 4, Application US/10363112  
 ; Publication No. US20040091964A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: THE AUSTRALIAN NATIONAL UNIVERSITY  
 ; APPLICANT: BOARD, PHILLIP  
 ; APPLICANT: HARRIS, MATTHEW  
 ; TITLE OF INVENTION: MODIFIED PROTEINS, ISOLATED NOVEL PEPTIDES, AND USES THEREOF  
 ; FILE REFERENCE: 007643-0302189  
 ; CURRENT APPLICATION NUMBER: US/10/363,112  
 ; CURRENT FILING DATE: 2003-11-03  
 ; PRIOR APPLICATION NUMBER: PCT/AU01/01093  
 ; PRIOR FILING DATE: 2001-08-31  
 ; PRIOR APPLICATION NUMBER: 60/229,663  
 ; PRIOR FILING DATE: 2000-08-31  
 ; NUMBER OF SEQ ID NOS: 62





THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:19 ; Search time 138.282 Seconds  
(without alignments)  
5389.603 Million cell updates/sec

Title: US-10-665-283-1

Perfect score: 9903

Sequence: 1 MALRGFCADGSDPLDMNV.....RSVAVAKPKFSPDLS 1927

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001s.\*
- 5: Geneseq2002s.\*
- 6: Geneseq2003s.\*
- 7: Geneseq2003bs.\*
- 8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7860	79.4	1531	2	Aaw57486 Human MRP
2	7860	79.4	1531	2	Aaw74471 Human mul
3	7860	79.4	1531	2	Aaw99894 Human mul
4	7860	79.4	1531	3	Aay55799 Human mul
5	7860	79.4	1531	3	Aay78873 Multidrug
6	7860	79.4	1531	5	ABG61810 Prostate
7	7860	79.4	1531	6	ABM35012 Cancer ba
8	7860	79.4	1531	6	ADB20865 MRP1 base
9	7860	79.4	1531	7	ADB87954 Human UGT
10	7860	79.4	1531	7	ADB96937 Human MDR
11	7860	79.4	1531	7	ADB92128 Human MDR
12	7860	79.4	1531	7	ADD44764 Human Pro
13	7860	79.4	1531	8	ADN97111 MRP1 prot
14	7856	79.3	1530	7	ADN95929 Human BEC
15	7849	79.3	1531	2	Aar54928 Human MRP
16	7849	79.3	1531	2	Aar93153 Multidrug
17	7849	79.3	1531	2	Aaw57485 Human mul
18	7849	79.3	1531	2	Aaw74470 Human mul
19	7849	79.3	1531	2	Aaw99893 Human mul
20	7849	79.3	1531	3	Aay55798 Human mul
21	7849	79.3	1531	3	Aay78872 Human mul
22	7849	79.3	1531	3	ABM3582 Multidrug
23	7810	76.8	1489	2	AAR96952 Multidrug
24	7587.5	76.6	1482	8	ADN03902 Antipsori
25	7587.5	76.6	1482	8	ADP23422 PRO polyp

26	7572	76.5	1481	5	ABP52108 Homo sapi
27	7560.5	76.3	1482	2	AAR96953 Multi-dru
28	7500.5	75.7	1472	2	AAR96954 Multi-dru
29	7230.5	73.0	1530	7	ADF56611 Bovine MR
30	7194	72.6	1417	2	AAR96955 Multi-dru
31	7037	71.1	1388	7	ADJ69912 Human hea
32	7002.5	70.7	1528	2	AAW57487 Murine mu
33	7002.5	70.7	1528	2	AAW74472 Mouse mul
34	7002.5	70.7	1528	2	AAW99895 Mouse mul
35	7002.5	70.7	1528	3	AAV55800 Murine mu
36	7002.5	70.7	1528	3	AAV78874 Murine mu
37	6916	69.8	1372	2	AAR98211 Multi-dru
38	6911	69.8	1345	2	AAR96956 Multi-dru
39	4491.5	45.4	1527	5	AAU91309 Human pro
40	4491.5	45.4	1527	5	ABP52110 Homo sapi
41	4491.5	45.4	1527	6	ABR58639 Human can
42	4491.5	45.4	1527	7	ADP54387 Human pro
43	4491.5	45.4	1527	7	ADN40024 Cancer/an
44	4491.5	45.4	1527	8	ADP24088 PRO polyp
45	4487.5	45.3	1527	2	AAV43543 A human M

#### ALIGNMENTS

##### RESULT 1

AAW57486

ID AAW57486 standard; protein; 1531 AA.

XX AAW57486;

DT 14-AUG-1998 (first entry)

DE Human MRP variant ltPgpA (Lei/PgpA).

XX Multidrug resistance-associated protein; MRP; tumour; human; variant;  
KW multidrug resistance; MDR; leishmania P-glycoprotein; ltPgpA; Lei/PgpA.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 685

FT /label= L685S

FT /note= "wild-type Leu is replaced by Ser"

FT Misc-difference 1282

FT /label= R1282A

FT /note= "wild-type Arg is replaced by Ala"

XX US576680-A.

XX 16-JUN-1998.

XX 05-JUN-1995; 95US-00463092.

XX 27-OCT-1992; 92US-00966923.

XX 08-MAR-1993; 93US-00029340.

XX 26-OCT-1993; 93US-00141893.

XX 20-MAR-1995; 95US-00407207.

XX (TOOH ) UNIV QUEENS KINGSTON.

XX Cole SP, Deeley RG;

XX WPI; 1998-361687/31.

XX N-FSDB; AAV31498.

DNA encoding protein associated with multi-drug resistance - useful for  
as probe for identifying multi-drug resistant tumour cells.  
Claim 1; Col 67-78; 82pp; English.

XX This represents a variant of the human multidrug resistance-associated  
protein (MRP). This natural variant is a leishmania P-glycoprotein related

CC molecule ltpgpa (Lei/Pgpa). The human and murine MRP nucleic acid  
 CC molecules can be used as probes for identifying multidrug resistant  
 CC tumour cells by hybridisation to mRNA from tumour cells. The antisense  
 CC nucleic acid can be used to reverse multidrug resistance (MDR). A  
 CC recombinant expression vector containing the MRP nucleic acid molecules  
 CC operatively linked to at least one regulatory sequence can be used to  
 CC transform a host cell to produce a recombinant MDR-associated protein  
 XX  
 SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 2; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDNNTWNTSNPDFTKCFQNTVLVWPCFYLWACFPFYFLYLSRH 60  
 DB 1 MALRGFCSADGSDPLMDNNTWNTSNPDFTKCFQNTVLVWPCFYLWACFPFYFLYLSRH 60

QY 61 DRGYIQMTPLNKTALGFLWIVCWADLFYFWSRSGIFLAPVPLVSPULLGTTLLA 120  
 DB 61 DRGYIQMTPLNKTALGFLWIVCWADLFYFWSRSGIFLAPVPLVSPULLGTTLLA 120

QY 121 TELIOLERRKGVSIGIMTLFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TELIOLERRKGVSIGIMTLFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180

QY 181 LLLIOLVLSCFSDRPLSETIHDNPNCPSSASFLSRTITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LLLIOLVLSCFSDRPLSETIHDNPNCPSSASFLSRTITFWITGLIVRGYRQPLEGSD 240

QY 241 LWSLNKEDTSEQVPLVQNMKCEKAKTRKQPVVYSSKDPAPKESKVDANEVEAL 300  
 DB 241 LWSLNKEDTSEQVPLVQNMKCEKAKTRKQPVVYSSKDPAPKESKVDANEVEAL 300

QY 301 IVKSPQKWNPSLFKVLKTYKTFPGYFLMFFFKAIHDLMMFSGPQILKLIKFVNDTKAPD 360  
 DB 301 IVKSPQKWNPSLFKVLKTYKTFPGYFLMFFFKAIHDLMMFSGPQILKLIKFVNDTKAPD 360

QY 361 WQGYEYTVLLFVACLOTILVHOYFHI CFVSGMRIKTAIVGAVYKALVITNSARKSTV 420  
 DB 361 WQGYEYTVLLFVACLOTILVHOYFHI CFVSGMRIKTAIVGAVYKALVITNSARKSTV 420

QY 421 GEIVNLSVDAQRFMDLATYINMWSAPLOVILALYLWNLGSPVLGAVVWMLVWPN 480  
 DB 421 GEIVNLSVDAQRFMDLATYINMWSAPLOVILALYLWNLGSPVLGAVVWMLVWPN 480

QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIHQBELKVLK 540  
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIHQBELKVLK 540

QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLAFNLRLPPLNLP 600  
 DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLAFNLRLPPLNLP 600

QY 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERPVPKGGGINSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERPVPKGGGINSITVRNATFTWARSDDPT 660

QY 661 LINGITFSIPEGALVAVGVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
 DB 661 LINGITFSIPEGALVAVGVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720

QY 721 SLRENILFGCOLLEPYYSVTOACALLPDLLETLPESGDRTEIGKGVNLSSGQKQVSLAR 780  
 DB 721 SLRENILFGCOLLEPYYSVTOACALLPDLLETLPESGDRTEIGKGVNLSSGQKQVSLAR 780

QY 781 AVYSNADILYLFDDPLSADVAHVGHIFENVIKPGKMLKNKTRILVTHSMSYLPQVDVLIIV 840  
 DB 781 AVYSNADILYLFDDPLSADVAHVGHIFENVIKPGKMLKNKTRILVTHSMSYLPQVDVLIIV 840

QY 841 MSGGKISEMGVQYQELLARDGAFAPFLRYTASTEQSDAEENGVTGVSQPGKEAKOMENGM 900  
 DB 841 MSGGKISEMGVQYQELLARDGAFAPFLRYTASTEQSDAEENGVTGVSQPGKEAKOMENGM 900

QY 901 LVTDAGKQLOLQSSSSSYSGDISRHHNSTABEQAKAEETWKLMEADKAQTQGVKL 960  
 DB 901 LVTDAGKQLOLQSSSSSYSGDISRHHNSTABEQAKAEETWKLMEADKAQTQGVKL 960

QY 961 SVYDYMKAIGLFTISFLIFLPMCNHVSALASNYWLSLWTDPIVNGTQHTKVLRSVYG 1020  
 DB 961 SVYDYMKAIGLFTISFLIFLPMCNHVSALASNYWLSLWTDPIVNGTQHTKVLRSVYG 1020

QY 1021 ALGISQGIAGVFGYSMAVISGIGILASRCLHVDLLHSILSPMSFFERTPSGNLVNRFSKEL 1080  
 DB 1021 ALGISQGIAGVFGYSMAVISGIGILASRCLHVDLLHSILSPMSFFERTPSGNLVNRFSKEL 1080

QY 1081 DTVDMSIPEVIMPMFGLFNIVIGACIVILLATPATAAIIIPPLGLIYFFVQRPYVASSQL 1140  
 DB 1081 DTVDMSIPEVIMPMFGLFNIVIGACIVILLATPATAAIIIPPLGLIYFFVQRPYVASSQL 1140

QY 1141 KRLESVSRSRSPVYSHFNELTLLGVSVIRAFEEQERFIHQSDLVKVDENOKAYYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSRSPVYSHFNELTLLGVSVIRAFEEQERFIHQSDLVKVDENOKAYYPSIVANRWLA 1200

QY 1201 VRLECVGNCIYVLPALFAVIRSHSLSAGLVGLSVYSLSQVTTYLNWLVRMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIYVLPALFAVIRSHSLSAGLVGLSVYSLSQVTTYLNWLVRMSSEMETNIVA 1260

QY 1261 VERLKEYSETEKEAPWQIQETAPPSSWPQVGRVFRNYCLYREDLDFVLRIHNVITNGG 1320  
 DB 1261 VERLKEYSETEKEAPWQIQETAPPSSWPQVGRVFRNYCLYREDLDFVLRIHNVITNGG 1320

QY 1321 EKVGIVGRTGAGKSLTLGLFRINESAGEIIDIIGINIAKIGLHDLRPKTIIPQDPVLF 1380  
 DB 1321 EKVGIVGRTGAGKSLTLGLFRINESAGEIIDIIGINIAKIGLHDLRPKTIIPQDPVLF 1380

QY 1381 SGSLRMNLDPFSPQYSDEEVTSLSLAHLKDFVSALPKDLHECAEGGNSLVGQRLVCL 1440  
 DB 1381 SGSLRMNLDPFSPQYSDEEVTSLSLAHLKDFVSALPKDLHECAEGGNSLVGQRLVCL 1440

QY 1441 ABALLRKTILVLDDEATAAVDLEDDLIQSTIRQTQFEDCTVLTIAHRLNTIMDYTRVIL 1500  
 DB 1441 ABALLRKTILVLDDEATAAVDLEDDLIQSTIRQTQFEDCTVLTIAHRLNTIMDYTRVIL 1500

QY 1501 DKGIQIEYGAPSDLLQQRGLFYSMADAGLV 1531  
 DB 1501 DKGIQIEYGAPSDLLQQRGLFYSMADAGLV 1531

RESULT 2  
 AAW74471  
 ID AAW74471 standard; protein, 1531 AA.  
 XX AAW74471;  
 AC AAW74471;  
 XX 18-MAY-1999 (first entry)  
 DT Human multidrug resistance-associated protein variant.  
 XX Multidrug resistance-associated protein; MDR; human; diagnosis;  
 KW MDR tumour cell identification; cancer therapy.  
 XX Homo sapiens.  
 XX US5882875-A.  
 XX 16-MAR-1999.  
 XX 05-JUN-1995; 95US-00462109.  
 XX 27-OCT-1992; 92US-00966923.  
 PR 08-MAR-1993; 93US-00029340.  
 PR 26-OCT-1993; 93US-00141893.  
 PR 20-MAR-1995; 95US-00407207.  
 XX (TOOH ) UNIV QUEENS KINGSTON.  
 PA





[illegible]

Db 1321 ERVIGVIRGTAGKSSLTGLPRINSAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 Qy 1381 SGLRMNLDPPFSQYSDDEEVTTSLELAHLKDFVSALPDKLDHCEAGGNNLSVGQQLVCL 1440  
 Db 1381 SGLRMNLDPPFSQYSDDEEVTTSLELAHLKDFVSALPDKLDHCEAGGNNLSVGQQLVCL 1440  
 Qy 1441 ARALLRKTKILVLDATAVDLETDLIQSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTKILVLDATAVDLETDLIQSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Qy 1501 DKGEIQEYCAPSDLLQORGLFYSMADAGLV 1531  
 Db 1501 DKGEIQEYCAPSDLLQORGLFYSMADAGLV 1531

RESULT 4

AA55799  
 ID AA55799 standard; protein; 1531 AA.  
 AC AA55799;  
 XX 28-FEB-2000 (first entry)  
 XX Human multidrug resistance-associated protein (MRP) variant.  
 DE Chemosensitizer; multidrug resistance-associated protein; MRP; human;  
 KW therapeutic agent; P-glycoprotein-mediated multidrug resistance; lung;  
 KW cancer; variant.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FH Misc-difference 685 /label= L685S  
 FT /note= "wild-type Leu is replaced by Ser"  
 FT Misc-difference 1282 /label= R1282A  
 FT /note= "wild-type Arg is replaced by Ala"  
 XX US6001563-A.  
 XX 14-DEC-1999.  
 XX 05-JUN-1995; 95US-00463179.  
 XX 27-OCT-1992; 92US-00966923.  
 PR 08-MAR-1993; 93US-00029340.  
 PR 26-OCT-1993; 93US-00141893.  
 PR 20-MAR-1995; 95US-00407207.  
 XX (TOOH ) UNIV QUEENS KINGSTON.  
 XX Cole SP, Deeley RG;  
 XX WPI; 2000-061877/05.  
 DR N-PSDB; AA239556.  
 XX Identification of chemosensitizers useful for treating cancer, using  
 FT nucleic acids encoding multidrug resistance-associated protein.  
 XX Claim 3; Col 69-80; 77pp; English.  
 XX The invention provides a method for identifying a substance which is a  
 CC chemosensitizer that comprises, contacting a cell transfected with  
 CC nucleic acid encoding multidrug resistance-associated protein (MRP) with  
 CC a therapeutic agent in vitro. The method is useful for identifying  
 CC chemosensitizers which may then be used to treat cancer (especially lung  
 CC cancer). The method allows the identification of chemosensitizers which  
 CC do not reverse P-glycoprotein-mediated multidrug resistance. The present  
 CC sequence represents a human MRP variant  
 XX Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 3; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MALRGFCSDGSDPLNDWNTWNTSNPDFTKCFQNTLVWVPCFVLMACFFPFYLYLSRH 60  
 Db 1 MALRGFCSDGSDPLNDWNTWNTSNPDFTKCFQNTLVWVPCFVLMACFFPFYLYLSRH 60  
 Qy 61 DRGIQMTPLNKTALGFLWVCWADLFYSFWERSRGIFLAPVFLVSPILLGTTLLA 120  
 Db 61 DRGIQMTPLNKTALGFLWVCWADLFYSFWERSRGIFLAPVFLVSPILLGTTLLA 120  
 Qy 121 TFLQLERRKGVQSSGIMLTFWLVALVCAILRSKIMTALKEDAQVDLFRDITFYVVS 180  
 Db 121 TFLQLERRKGVQSSGIMLTFWLVALVCAILRSKIMTALKEDAQVDLFRDITFYVVS 180  
 Qy 181 LLLIQLVLSCFSDRSPLESETIHDNPNCPESASFLSRIITFWITGLIVRGYRQPLEGSD 240  
 Db 181 LLLIQLVLSCFSDRSPLESETIHDNPNCPESASFLSRIITFWITGLIVRGYRQPLEGSD 240  
 Qy 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKESKVDANEVEAL 300  
 Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKESKVDANEVEAL 300  
 Qy 301 IVKSPQKEMNPSLFKVLKTYKTFGPYPLMSFFPKAIHDLMMFSGPQLKLLIKPVNDTKAPD 360  
 Db 301 IVKSPQKEMNPSLFKVLKTYKTFGPYPLMSFFPKAIHDLMMFSGPQLKLLIKPVNDTKAPD 360  
 Qy 361 WQGYFYTVLLFVTAQLQTLVHLHOYFHICFVSGMRIKTAIVGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTAQLQTLVHLHOYFHICFVSGMRIKTAIVGAVYRKALVITNSARKSSTV 420  
 Qy 421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLLNLGSPVLAGVAVMLMVPVN 480  
 Db 421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLLNLGSPVLAGVAVMLMVPVN 480  
 Qy 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 Qy 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAOTAFVSLALFNILRFPNLTLP 600  
 Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAOTAFVSLALFNILRFPNLTLP 600  
 Qy 601 MVISIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660  
 Db 601 MVISIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660  
 Qy 661 LNGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720  
 Db 661 LNGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720  
 Qy 721 SIRENIFGCOLLEPYRSVIOACALLPDLEILPSGDRTEIGEKGVNLSSGGOKQVSLAR 780  
 Db 721 SIRENIFGCOLLEPYRSVIOACALLPDLEILPSGDRTEIGEKGVNLSSGGOKQVSLAR 780  
 Qy 781 AVYSNADILYLDPLSADVAHVGHKIFENVIGPKGMLKNKTRILVTHSHSYLPQVDVIV 840  
 Db 781 AVYSNADILYLDPLSADVAHVGHKIFENVIGPKGMLKNKTRILVTHSHSYLPQVDVIV 840  
 Qy 841 MSGGKISMGYSQVELLARDGAPAEFLRTYASTEQEQDABENGVTGVSGGKAKQENGM 900  
 Db 841 MSGGKISMGYSQVELLARDGAPAEFLRTYASTEQEQDABENGVTGVSGGKAKQENGM 900  
 Qy 901 LVTSAGLQORQLSSSSSYSGDISRHNSIAELQAKAKKEETWKLMEADKAQTQGVKL 960  
 Db 901 LVTSAGLQORQLSSSSSYSGDISRHNSIAELQAKAKKEETWKLMEADKAQTQGVKL 960  
 Qy 961 SVYDYMKAIGLFTISFLSIFLFCMCHVNSALASNYLWLTDDPIVNGTQEHKTVRLSVYG 1020  
 Db 961 SVYDYMKAIGLFTISFLSIFLFCMCHVNSALASNYLWLTDDPIVNGTQEHKTVRLSVYG 1020  
 Qy 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHLSILRSPMSFFERTPSGNLVNRFSEL 1080

Db 1021 ALGISQIAVFGYSMAVSGIGILASRCLHVDLHLSLRSPMSFFERTPSGNLVNRFSEL 1080  
 Qy 1081 DTVDMSIPEVIMKFMGSLFNWIGACIVILLATPIAAIIPIPLGLIYFFVQRPYVASSRQL 1140  
 Db 1081 DTVDMSIPEVIMKFMGSLFNWIGACIVILLATPIAAIIPIPLGLIYFFVQRPYVASSRQL 1140  
 Qy 1141 KELESVSRSPVYSHNETLLGVSVIRAEPEQERFHQSDLVKVDENOKAYPSIVANRWLA 1200  
 Db 1141 KELESVSRSPVYSHNETLLGVSVIRAEPEQERFHQSDLVKVDENOKAYPSIVANRWLA 1200  
 Qy 1201 VRLCEVGCNIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTLYLNLVMSSEMETNIVA 1260  
 Db 1201 VRLCEVGCNIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTLYLNLVMSSEMETNIVA 1260  
 Qy 1261 VERLKEYSETEKAPWQIETAPPSWPQVGRVFNRYCLRVRELDLDFVLRHINVTINGG 1320  
 Db 1261 VERLKEYSETEKAPWQIETAPPSWPQVGRVFNRYCLRVRELDLDFVLRHINVTINGG 1320  
 Qy 1321 EKVGIVGRTGAKSSLTTLGLPRINSAEGEIIIDGINIAKIGLHDLRFKTIIPDDPVLF 1380  
 Db 1321 EKVGIVGRTGAKSSLTTLGLPRINSAEGEIIIDGINIAKIGLHDLRFKTIIPDDPVLF 1380  
 Qy 1381 SGSLRMNLDPFQYSDEEYVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 Db 1381 SGSLRMNLDPFQYSDEEYVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 Qy 1441 ABALLRKTKILVLDATAAVDLETDLLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ABALLRKTKILVLDATAAVDLETDLLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Qy 1501 DKGEIQEYGAQPSDLLQQRGLFYSMAKDAAGLV 1531  
 Db 1501 DKGEIQEYGAQPSDLLQQRGLFYSMAKDAAGLV 1531

RESULT 5

AAV78873  
 ID AAV78873 standard; protein; 1531 AA.  
 XX AAV78873;  
 AC AAV78873;  
 XX  
 DT 19-MAY-2000 (first entry)  
 XX  
 DE Multidrug resistance protein (MRP) natural variant amino acid sequence.  
 XX  
 KW Multidrug resistance protein; MRP; human; anthracycline; Vinca alkaloid;  
 XW epipodophyllotoxin; cancer; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 FN US6025473-A.  
 PD 15-FEB-2000.  
 XX  
 PF 05-JUN-1995; 95US-00461384.  
 XX  
 PR 27-OCT-1992; 92US-00966923.  
 PR 08-MAR-1993; 93US-00029340.  
 PR 26-OCT-1993; 93US-00141893.  
 PR 20-MAR-1995; 95US-00407207.  
 XX  
 PA (TOOH ) UNIV QUEBENS KINGSTON.  
 XX  
 PI Cole SPC, Deeley RG;  
 XX  
 DR WPI; 2000-181838/16.  
 DR N-PSDB; AA290193.  
 XX  
 PT Isolated protein conferring multidrug resistance, to at least two drugs  
 PT selected from anthracyclines, epipodophyllotoxins and Vinca alkaloids, on  
 PT a drug sensitive mammalian cell.  
 XX

PS Claim 10; Col 79-88; 78pp; English.  
 XX  
 CC This sequence represents a human multidrug resistance protein (MRP)  
 CC natural variant amino acid sequence. The human MRP confers multidrug  
 CC resistance, including resistance to at least two drugs selected from  
 CC anthracyclines, epipodophyllotoxins and Vinca alkaloids, on a drug  
 CC sensitive mammalian cell, when the protein is expressed in the cell. The  
 CC multidrug resistance is not substantially reversed by chemosensitizers  
 CC which reverse P-glycoprotein-mediated multidrug resistance. The MRP  
 CC protein sequence can be used to generate antibodies against MRP. The MRP  
 CC protein and nucleotide sequences can be used in compositions which are  
 CC used to treat patients with tumours displaying multidrug resistance. The  
 CC compositions and methods of the invention can be used particularly to  
 CC treat breast cancer, leukaemias, fibrosarcomas, cervical cancer, gliomas,  
 CC thymomas, neuroblastomas, and lung cancer. Antibodies directed against  
 CC MRP can be used to inhibit the multidrug resistance of a multidrug  
 CC resistant cell  
 XX  
 SQ Sequence 1531 AA;  
 Query Match 79.4%; Score 7860; DB 3; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MALRGFCSADGSDPLWDNVTWNTSNPDTKCFQNTVLVWVPCFVLWACFPFFLYLSRH 60  
 Db 1 MALRGFCSADGSDPLWDNVTWNTSNPDTKCFQNTVLVWVPCFVLWACFPFFLYLSRH 60  
 Qy 61 DRGYQMTPLNKTALGFLWIVCWADLFYSFWSRGIFLAPVFLVSPITLLGTTLLA 120  
 Db 61 DRGYQMTPLNKTALGFLWIVCWADLFYSFWSRGIFLAPVFLVSPITLLGTTLLA 120  
 Qy 121 TFLIQLERRKGVQSSGIMLTFWLVNLCALAILSKIMTALKEDAQVDLPDIFYYVYS 180  
 Db 121 TFLIQLERRKGVQSSGIMLTFWLVNLCALAILSKIMTALKEDAQVDLPDIFYYVYS 180  
 Qy 181 LLLIQLVLSFSDRSPLFSETHDPNCPSSASFLSRITFWITGLIVRGYRPLEGSD 240  
 Db 181 LLLIQLVLSFSDRSPLFSETHDPNCPSSASFLSRITFWITGLIVRGYRPLEGSD 240  
 Qy 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVKVYSSKDDPAQPKSSKVDANEVEAL 300  
 Db 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVKVYSSKDDPAQPKSSKVDANEVEAL 300  
 Qy 301 IVKSPQKEWNPSPKVLKYKTFPGPYFLMSFFPKAIDHLMFSGPOLLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKEWNPSPKVLKYKTFPGPYFLMSFFPKAIDHLMFSGPOLLLIKFVNDTKAPD 360  
 Qy 361 WQGYFYTVLLFVTACLTQLVLHQLVHFHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTACLTQLVLHQLVHFHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
 Qy 421 GEIVNLSVDAQRPMDLATYINMIWSAPLQVILALYLLNLGSPSVLAGVAMVLMVFN 480  
 Db 421 GEIVNLSVDAQRPMDLATYINMIWSAPLQVILALYLLNLGSPSVLAGVAMVLMVFN 480  
 Qy 481 AVMAKTKTYQVAHMKSKDNRIKLWNETLNGIKVLKYAWELAFKDKVLAIHQBELKVLK 540  
 Db 481 AVMAKTKTYQVAHMKSKDNRIKLWNETLNGIKVLKYAWELAFKDKVLAIHQBELKVLK 540  
 Qy 541 KSAVLSAVGTFTWVCTPFLVALCTFAVVYVITDENNILDQAQTAFAVSLALFNILRFLNLP 600  
 Db 541 KSAVLSAVGTFTWVCTPFLVALCTFAVVYVITDENNILDQAQTAFAVSLALFNILRFLNLP 600  
 Qy 601 MVISSIVQASVSLKRLRIFLSHELEPDSITERRPVKGGGTNSITVRNATTTWARSDDPT 660  
 Db 601 MVISSIVQASVSLKRLRIFLSHELEPDSITERRPVKGGGTNSITVRNATTTWARSDDPT 660  
 Qy 661 LNGTHFTSPEGALVAVVGVCCKSSLLSALLAEMDKVEGHVAKGVSVAVYVQQAQWIND 720  
 Db 661 LNGTHFTSPEGALVAVVGVCCKSSLLSALLAEMDKVEGHVAKGVSVAVYVQQAQWIND 720  
 Qy 721 SLRENILFGCQLEBPYRYSVIQACALLPDLFELPSGDRTEIGKGVNLSGQKQKRVSLAR 780

Db 721 SURENIFGCOLEBEEYRSVIOACALLPDLLEIPSGDRTEIGEGKGNLSGGOKQVSLAR 780  
Qy 781 AVYSNADIYLPDPLSANDVHGVKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVLIIV 840  
Db 781 AVYSNADIYLPDPLSANDVHGVKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVLIIV 840  
Qy 841 MSGKISMGSGYQELLARDGAPAEFLRTYASTEQDABEENGVTGVSFGPGKEAKOMNGM 900  
Db 841 MSGKISMGSGYQELLARDGAPAEFLRTYASTEQDABEENGVTGVSFGPGKEAKOMNGM 900  
Qy 901 LVTSAGKOLQRLSSSSSYSGDISRHNSHSTAELOKAEAKKEETWKLMEADKAQTGVKL 960  
Db 901 LVTSAGKOLQRLSSSSSYSGDISRHNSHSTAELOKAEAKKEETWKLMEADKAQTGVKL 960  
Qy 961 SVYDYMKAIGLIFSLFIFCMCHVNSALASNYLSLWTDPIVNGTQETHKVLRSYVG 1020  
Db 961 SVYDYMKAIGLIFSLFIFCMCHVNSALASNYLSLWTDPIVNGTQETHKVLRSYVG 1020  
Qy 1021 ALGISQGIAGVGYMAVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSEL 1080  
Db 1021 ALGISQGIAGVGYMAVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSEL 1080  
Qy 1081 DTVDNMIPEVIMKFMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQFPYVASSRQL 1140  
Db 1081 DTVDNMIPEVIMKFMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQFPYVASSRQL 1140  
Qy 1141 KRLESVSRPVYSHNETLLGVSVTRAEEQERFIHQSDLVKDNOKAYPSIVANRWLA 1200  
Db 1141 KRLESVSRPVYSHNETLLGVSVTRAEEQERFIHQSDLVKDNOKAYPSIVANRWLA 1200  
Qy 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
Qy 1261 VERLKEYSETEKAPWQIQTAPPSPQVGRVPEPRNYCLRYREDLDFVLRHINVTINGG 1320  
Db 1261 VERLKEYSETEKAPWQIQTAPPSPQVGRVPEPRNYCLRYREDLDFVLRHINVTINGG 1320  
Qy 1321 EKVGIIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIPODPVLF 1380  
Db 1321 EKVGIIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIPODPVLF 1380  
Qy 1381 SGSLRMNLDPFQSYDEEVTWTSLELAHLKDFVSALPDKLDHECARGENLSVGQRLVCL 1440  
Db 1381 SGSLRMNLDPFQSYDEEVTWTSLELAHLKDFVSALPDKLDHECARGENLSVGQRLVCL 1440  
Qy 1441 APALLRKTKILVLDATAAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Db 1441 APALLRKTKILVLDATAAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Qy 1501 DKGEIOEYCAPSDLLQORGLFYSMADKAGLV 1531  
Db 1501 DKGEIOEYCAPSDLLQORGLFYSMADKAGLV 1531

RESULT 6  
ABG61810

ID ABG61810 standard; protein; 1531 AA.

XX AC ABG61810;

XX DT 15-AUG-2002 (first entry)

XX DE Prostate cancer-associated protein #11.

XX KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

XX OS Mammalia.

XX FN WO200202068-A2.

XX PD 18-APR-2002.

XX 12-OCT-2001; 2001WO-US032045.  
XX 13-OCT-2000; 2000US-00687576.  
PR 08-DEC-2000; 2000US-00733288.  
PR 08-DEC-2000; 2000US-00733742.  
PR 24-JAN-2001; 2001US-0263957P.  
PR 16-MAR-2001; 2001US-0276791P.  
PR 06-APR-2001; 2001US-0276888P.  
PR 24-APR-2001; 2001US-0286214P.  
PR 30-APR-2001; 2001US-00847046.  
PR 04-MAY-2001; 2001US-0288589P.  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;  
XX WPI; 2002-471335/50.  
XX N-PSDB; ABK92125.  
DR Detecting a prostate cancer-associated transcript in a cell in a patient,  
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,  
PT by determining if prostate cancer-associated genes are expressed in a  
PT prostate tissue.  
PS Claim 27; Page 309; 436pp; English.  
XX The present invention relates to methods of detecting a prostate cancer-  
CC associated transcript in a cell from a patient. The method comprises  
CC contacting a biological sample from the patient with prostate cancer-  
CC associated polynucleotides (designated PC genes) that selectively  
CC hybridise to a sequence that is at least 80% identical to them. The  
CC prostate cancer-associated polynucleotide sequences are differentially  
CC expressed in prostate tumour tissue or in prostate cancer and are derived  
CC from the tissues of various organisms such as humans or other mammals  
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for  
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-  
CC associated genes are useful for diagnosing or treating prostate cancer,  
CC as well as for identifying modulators of prostate cancer or agents that  
CC inhibit prostate cancer. The nucleic acid sequences are particularly  
CC useful in gene therapy, as a vaccine or in antisense applications.  
CC ABG61800-ABG61944 represent prostate cancer-associated proteins  
XX Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 5; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLMDWNTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60  
Db 1 MALRGFCSADGSDPLMDWNTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60  
Qy 61 DRGYQMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPDLLGTTLLA 120  
Db 61 DRGYQMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPDLLGTTLLA 120  
Qy 121 TFLQLERRKGVQSSGIMLTFLWALVCAILSKIMTALKEDAQVDFLFRDITVWVYFS 180  
Db 121 TFLQLERRKGVQSSGIMLTFLWALVCAILSKIMTALKEDAQVDFLFRDITVWVYFS 180  
Qy 181 LLLQLVLSCFSDRSPLESETIHDNPNCPESASFLSRITFWITGLIVRGVQPLEGSD 240  
Db 181 LLLQLVLSCFSDRSPLESETIHDNPNCPESASFLSRITFWITGLIVRGVQPLEGSD 240  
Qy 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKQPKVYVSSKDPAPQPKSSKVDANEVEAL 300  
Db 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKQPKVYVSSKDPAPQPKSSKVDANEVEAL 300  
Qy 301 IVKSPQKEWNPFLFKVLYKTGPIFLMSFFFKAIHDLWMFSGPQLKLLIKFVNDTKAPD 360  
Db 301 IVKSPQKEWNPFLFKVLYKTGPIFLMSFFFKAIHDLWMFSGPQLKLLIKFVNDTKAPD 360

QY 361 WQGYEYTVLLFVTVACQTLVLHVOYPHICFVSGMRKTAIVIGAVYKALVITNSARKSTV 420  
 Db 361 WQGYEYTVLLFVTVACQTLVLHVOYPHICFVSGMRKTAIVIGAVYKALVITNSARKSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLWNLGSPVLGAVVWVLMVFN 480  
 Db 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLWNLGSPVLGAVVWVLMVFN 480  
 QY 481 AVMAKTKTYQVAHMKSDNRKIKMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 Db 481 AVMAKTKTYQVAHMKSDNRKIKMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFAFSLALFNILRPLNLTLP 600  
 Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFAFSLALFNILRPLNLTLP 600  
 QY 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 Db 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 QY 661 LINGITFSIPEGALVAVGVQGVGKSSLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
 Db 661 LINGITFSIPEGALVAVGVQGVGKSSLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
 QY 721 SLRENILFGCOLLEPYYSVIOACALLPDLEILPSGDRTEIGEGVNLSSGQKORVSLAR 780  
 Db 721 SLRENILFGCOLLEPYYSVIOACALLPDLEILPSGDRTEIGEGVNLSSGQKORVSLAR 780  
 QY 781 AVYSNADLYLFDPPLSAVDAHVGHKIFENVIGPKGMLKXKTRILVTHSMSYLPQVDVIIV 840  
 Db 781 AVYSNADLYLFDPPLSAVDAHVGHKIFENVIGPKGMLKXKTRILVTHSMSYLPQVDVIIV 840  
 QY 841 MSGGKISEMGYSQELLARDGAFAPLRTYASTEQDAEENGVTGVSQPGKEAKQNGM 900  
 Db 841 MSGGKISEMGYSQELLARDGAFAPLRTYASTEQDAEENGVTGVSQPGKEAKQNGM 900  
 QY 901 LVTDGAGLQORQLSSSSSYSGDISRHNSHTAELOKAEKEETWKLMEADKAOTGOVKL 960  
 Db 901 LVTDGAGLQORQLSSSSSYSGDISRHNSHTAELOKAEKEETWKLMEADKAOTGOVKL 960  
 QY 961 SVYDYMKAIGLFIISFLSIFLMCNHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 Db 961 SVYDYMKAIGLFIISFLSIFLMCNHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 QY 1021 ALGISQGIATVFGYSNAVSGIGILASRCLHVDLLHSILRSPSPFERTPSGNLVRFSKEL 1080  
 Db 1021 ALGISQGIATVFGYSNAVSGIGILASRCLHVDLLHSILRSPSPFERTPSGNLVRFSKEL 1080  
 QY 1081 DTVDMSIPEVIMKFMGSLFNVIGACIVILLATPTAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 Db 1081 DTVDMSIPEVIMKFMGSLFNVIGACIVILLATPTAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHFNETHLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 Db 1141 KRLESVSRSPVYSHFNETHLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVIRSRHLSAGLGLSVSYSLQVTTYLNLWLRMSSEMETNIVA 1260  
 Db 1201 VRLECVGNCIVLFAALFAVIRSRHLSAGLGLSVSYSLQVTTYLNLWLRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKEAPWQIQETAPSSWPQGRVFRNCLRYREDLDFVLARHINVTNGG 1320  
 Db 1261 VERLKEYSETEKEAPWQIQETAPSSWPQGRVFRNCLRYREDLDFVLARHINVTNGG 1320  
 QY 1321 EKVGVIGRTGAGKSSLTGLFRINESARGEIIDIINIAKIGLHDLRPFKTIIPQDPVLF 1380  
 Db 1321 EKVGVIGRTGAGKSSLTGLFRINESARGEIIDIINIAKIGLHDLRPFKTIIPQDPVLF 1380  
 QY 1381 SGLSLMNLDPQSQSDSEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQROLVCL 1440  
 Db 1381 SGLSLMNLDPQSQSDSEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQROLVCL 1440

QY 1441 ARALLRKTILVLDEATAAVIDLETTDLIQSTIRQFEDCTVLTIAHRLNTIMDYTRVILV 1500  
 Db 1441 ARALLRKTILVLDEATAAVIDLETTDLIQSTIRQFEDCTVLTIAHRLNTIMDYTRVILV 1500  
 QY 1501 DKGBIQEYGAPESDLLQQRGLFYFSMAKDAGLV 1531  
 Db 1501 DKGBIQEYGAPESDLLQQRGLFYFSMAKDAGLV 1531

RESULT 7  
 ABM35012  
 ID ABM35012 standard; protein; 1531 AA.  
 XX  
 AC ABM35012;  
 XX  
 DT 08-OCT-2003 (first entry)  
 XX  
 DE Cancer based on CYP3A5 related protein SEQ ID NO:678.  
 XX  
 KW Cancer; CYP3A5; irinotecan; pharmaceutical; malignant glioma;  
 KW cytochrome p450; subfamily IIIA; nifedipine oxidase; polypeptide 5;  
 KW cytosstatic.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2003013534-A2.  
 XX  
 PD 20-FEB-2003.  
 XX  
 PF 23-JUL-2002; 2002WO-EP008219.  
 XX  
 PR 23-JUL-2001; 2001EP-00117608.  
 PR 24-MAY-2002; 2002EP-00011710.  
 XX  
 PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
 XX  
 PI Heinrich G, Kerb R;  
 XX  
 DR WPT; 2003-268144/26.  
 XX  
 PT New use of irinotecan for preparation of compositions for treating cancer  
 PT in subject having genome with variant allele comprising cytochrome p450,  
 PT subfamily IIIA, polypeptide 5 polynucleotide, termed CYP3A5.  
 XX  
 PS Disclosure; SEQ ID NO 678; 86pp; English.  
 XX  
 CC The present invention describes the use of irinotecan (I) or its  
 CC derivative for the preparation of a pharmaceutical composition for  
 CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic  
 CC cancer, or malignant glioma in a subject having a genome with a variant  
 CC allele which comprises a cytochrome p450, subfamily IIIA (nifedipine  
 CC oxidase), polypeptide 5 (CYP3A5) polynucleotide (ii). (i) and (ii) have  
 CC cytosstatic activity. The therapeutic applications of (i) is improved,  
 CC since it is possible to individually treat a subject with an appropriate  
 CC dosage and/or an appropriate derivative of (i). Therefore, undesirable,  
 CC harmful or toxic effects are efficiently avoided. Necessary and  
 CC potentially harmful treatment of those subjects who do not respond to the  
 CC treatment with substances (nonresponders), as well as the development of  
 CC drug resistances due to suboptimal drug dosing can be avoided. AC62200  
 CC to AC62751 and ABM34912 to ABM35013 represent sequences used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 6; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAIRGFCASGSDPLWDMVNTWNTSNDFTKCFONTVLVWVPCFYLWACPFYLYLSRH 60  
 Db 1 MAIRGFCASGSDPLWDMVNTWNTSNDFTKCFONTVLVWVPCFYLWACPFYLYLSRH 60  
 QY 61 DRGYIQMTPLNKTALGFLLIWVCWADLFYFWSRSGIFLAPVFLVSLGTTLLA 120

Db 61 DRYIQMTPLNKTALGALLIVCWADLFYFWSRSGIFLAPVFLVSPITLLGITLLA 120  
 QY 121 TELIOLERRKGQSSGIMLTFWLVALVICALATLRSKIMTALKEDAQVDLFRDITFVYVS 180  
 Db 121 TELIOLERRKGQSSGIMLTFWLVALVICALATLRSKIMTALKEDAQVDLFRDITFVYVS 180  
 QY 181 LLLIQLVLSCFSDRSPFLSETIHDNPNCPSESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 Db 181 LLLIQLVLSCFSDRSPFLSETIHDNPNCPSESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVVYSSKDPQPKESKVDANEVEAL 300  
 Db 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVVYSSKDPQPKESKVDANEVEAL 300  
 QY 301 IVKSPQKWNPSLFLVLYKTPGYFLMSPFPAIHDLMFSGPQILKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKWNPSLFLVLYKTPGYFLMSPFPAIHDLMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTAQLTLVLHQYFHCIFVSGMRITKAVIGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTAQLTLVLHQYFHCIFVSGMRITKAVIGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINNIWSAPLOVILALYLLNLGSPVLGAVVNLMPVN 480  
 Db 421 GEIVNLSVDAQRFMDLATYINNIWSAPLOVILALYLLNLGSPVLGAVVNLMPVN 480  
 QY 481 AVAMKTKTYOAHWKSNDNRILKLNELINGIKVLKLYAWELAFKDKVLATROELKVLK 540  
 Db 481 AVAMKTKTYOAHWKSNDNRILKLNELINGIKVLKLYAWELAFKDKVLATROELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFAVSLAFNLILPFLNLP 600  
 Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFAVSLAFNLILPFLNLP 600  
 QY 601 WVISSIVQASVSLKRLRIPLSHEELEPOSIERRPVKDGGTNSITVRNATFTWASDPT 660  
 Db 601 WVISSIVQASVSLKRLRIPLSHEELEPOSIERRPVKDGGTNSITVRNATFTWASDPT 660  
 QY 661 LNGITFTSPEGALVAVGVQCGKSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIOND 720  
 Db 661 LNGITFTSPEGALVAVGVQCGKSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIOND 720  
 QY 721 SIRENIFGCOLLEPYRYSVIOACALLPDLPLPSGDRTEIGEGKGNLSGGQKQSVSLAR 780  
 Db 721 SIRENIFGCOLLEPYRYSVIOACALLPDLPLPSGDRTEIGEGKGNLSGGQKQSVSLAR 780  
 QY 781 AVYSNADYILFDDPLSAVDHVGKHI PENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 Db 781 AVYSNADYILFDDPLSAVDHVGKHI PENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 QY 841 MSGGKISEGSGYOELLARDGAFAELRTYASTEQEDAEENGVTGSGPKEAKOMENG 900  
 Db 841 MSGGKISEGSGYOELLARDGAFAELRTYASTEQEDAEENGVTGSGPKEAKOMENG 900  
 QY 901 LVTDSAGLQORQLSSSSYSYSDISRHNSSTAELQAKAEKTEWKLMEADKAQOGVKL 960  
 Db 901 LVTDSAGLQORQLSSSSYSYSDISRHNSSTAELQAKAEKTEWKLMEADKAQOGVKL 960  
 QY 961 SVYDYMKAIGLIFSLFIFPMCNHVSALASNYLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 Db 961 SVYDYMKAIGLIFSLFIFPMCNHVSALASNYLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 QY 1021 ALGISQGIATVFGYSMAVSTGGTILASRCLHVDLLHSILRSPMFFERTSGNIVNRFKEL 1080  
 Db 1021 ALGISQGIATVFGYSMAVSTGGTILASRCLHVDLLHSILRSPMFFERTSGNIVNRFKEL 1080  
 QY 1081 DTVDSMIEVIMKMGSLFNWIGACIVILLATPIAIIIPPLGLIYFFVQRYVASSRQL 1140  
 Db 1081 DTVDSMIEVIMKMGSLFNWIGACIVILLATPIAIIIPPLGLIYFFVQRYVASSRQL 1140  
 QY 1141 KRESVRSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200

Db 1141 KRESVRSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
 QY 1201 VRLECVGNICIVLFAALPAVIRSRHLSAGLVLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
 Db 1201 VRLECVGNICIVLFAALPAVIRSRHLSAGLVLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETKEAPWQIQETAPPSSWPQVGRVEFRNRYCLRYREDLDFVLRHINVTINGG 1320  
 Db 1261 VERLKEYSETKEAPWQIQETAPPSSWPQVGRVEFRNRYCLRYREDLDFVLRHINVTINGG 1320  
 QY 1321 EKVGVIGBTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 Db 1321 EKVGVIGBTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLSRMNLDPFSQYSDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 Db 1381 SGLSRMNLDPFSQYSDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 QY 1441 ARALLRKTKILVLEATAAVALDETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTKILVLEATAAVALDETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEYCAPSDLLIQORGLFYSMAKDAGLV 1531  
 Db 1501 DKGEIQEYCAPSDLLIQORGLFYSMAKDAGLV 1531

RESULT 8  
 ADB20865

ID ADB20865 standard; protein; 1531 AA.

XX ADB20865;

XX 20-NOV-2003 (first entry)

DE MRPI based cancer related protein SEQ ID NO:678.

KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;  
 KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;  
 KW variant allele; multidrug resistance protein 1; MRPI; cytostatic.

XX Unidentified.

XX WO2003013533-A2.

XX 20-FEB-2003.

XX 23-JUL-2002; 2002WO-EP008200.

XX 23-JUL-2001; 2001EP-00117608.

XX 24-MAY-2002; 2002EP-00011710.

XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

XX Heinrich G, Kerb R;

XX WPI; 2003-354397/33.

PT Use of irinotecan or its derivative for preparation of a pharmaceutical  
 PT composition for treating cancer in a subject having a genome with a  
 PT variant allele comprising a multidrug resistance protein 1  
 PT polynucleotide.

XX Disclosure; SEQ ID NO 678; 100pp; English.

CC The present invention describes a method for the use of irinotecan (I) or  
 CC its derivative for the preparation of a pharmaceutical composition for  
 CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic  
 CC cancer, or malignant glioma in a subject having a genome with a variant  
 CC allele which comprises a multidrug resistance protein 1 (MRPI)  
 CC polynucleotide (II). (I) has cytostatic activity. (II) or its derivative  
 CC can be used for the preparation of a pharmaceutical composition for  
 CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic



CC cancer, or malignant glioma in a subject, where the subject is a human  
 CC (preferably African or Asian) or a mouse. The present sequence represents  
 CC a sequence which is used in the exemplification of the present invention.  
 XX  
 SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 6; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MALRFGCSADGSDPDLWNNVTWNTGNPTKQFNTLVWVPCFYLWACFPFYLYLSRH	60
DB	1	MALRFGCSADGSDPDLWNNVTWNTGNPTKQFNTLVWVPCFYLWACFPFYLYLSRH	60
QY	61	DRGYIQMTPLNKTALGFLWICWADLFYSFWSRSGIFLAPVPLSPTLGLTTLLA	120
DB	61	DRGYIQMTPLNKTALGFLWICWADLFYSFWSRSGIFLAPVPLSPTLGLTTLLA	120
QY	121	TFLIQLERRKGQSSGIMLTFWLVALCALAIRSKIMTALKEDAQVDLFRDITFYVYFS	180
DB	121	TFLIQLERRKGQSSGIMLTFWLVALCALAIRSKIMTALKEDAQVDLFRDITFYVYFS	180
QY	181	LLLIQVLSCSDRSLPSETIHDNPNCPSSASFLSRITFWITGLIVRGVROPLEGSD	240
DB	181	LLLIQVLSCSDRSLPSETIHDNPNCPSSASFLSRITFWITGLIVRGVROPLEGSD	240
QY	241	LWSLNKEDTSEOVVPLVKNWKECAKTRKOPKVVYSSKOPAKPKSSKVDANEEVEAL	300
DB	241	LWSLNKEDTSEOVVPLVKNWKECAKTRKOPKVVYSSKOPAKPKSSKVDANEEVEAL	300
QY	301	IVKSPQKWNPSLFLKVLKTYKTPYFMSFFKAIHDLMMFSGPOLKLLIKFVNDTKAPD	360
DB	301	IVKSPQKWNPSLFLKVLKTYKTPYFMSFFKAIHDLMMFSGPOLKLLIKFVNDTKAPD	360
QY	361	WQGYFTYVLLFTVACLOTLVHQPFIHCFVSGMRITKAVIGAVRKALVITNSARKSTV	420
DB	361	WQGYFTYVLLFTVACLOTLVHQPFIHCFVSGMRITKAVIGAVRKALVITNSARKSTV	420
QY	421	GEIVNLSVDAQRFMDLATYINMISAPLOVILALYLWNLGPSVLGAVVWLWVFN	480
DB	421	GEIVNLSVDAQRFMDLATYINMISAPLOVILALYLWNLGPSVLGAVVWLWVFN	480
QY	481	AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK	540
DB	481	AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK	540
QY	541	KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDENNDLDAQTAFAVSLALFNLRPLNLP	600
DB	541	KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDENNDLDAQTAFAVSLALFNLRPLNLP	600
QY	601	MVISIVQASVSLKRLRIFLGSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT	660
DB	601	MVISIVQASVSLKRLRIFLGSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT	660
QY	661	LNGITFSIPEGALVAVVQGVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND	720
DB	661	LNGITFSIPEGALVAVVQGVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND	720
QY	721	SIRENIFGCOLLEBPYRSVTOACALLPDILPSGDRTEIGEGVNLSSGQKQVSLAR	780
DB	721	SIRENIFGCOLLEBPYRSVTOACALLPDILPSGDRTEIGEGVNLSSGQKQVSLAR	780
QY	781	AVYSNADITYLFDPLSADVAHVGHKIFENVIKPGKMLKNKTRILVTHSMSYLPQVDVLIIV	840
DB	781	AVYSNADITYLFDPLSADVAHVGHKIFENVIKPGKMLKNKTRILVTHSMSYLPQVDVLIIV	840
QY	841	MSGKISMGSGYQELLARDGAFABFLRTYASTEQDQAEENGVTGVSGPGKEAKOMENGM	900
DB	841	MSGKISMGSGYQELLARDGAFABFLRTYASTEQDQAEENGVTGVSGPGKEAKOMENGM	900
QY	901	LVTDSAGQOLORLSSSSSSSGDLSRHNSHNSAELOKAEAKETWKLMEADKATQGVKL	960
DB	901	LVTDSAGQOLORLSSSSSSSGDLSRHNSHNSAELOKAEAKETWKLMEADKATQGVKL	960

QY	961	SVTYDMKAIAGLFTSFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTOEHTKVPLSVYG	1020
DB	961	SVTYDMKAIAGLFTSFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTOEHTKVPLSVYG	1020
QY	1021	ALGISQGIAGVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL	1080
DB	1021	ALGISQGIAGVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL	1080
QY	1081	DTVDSMIDPEVTKMPGSLFNFGVIGACIVILLATPAAIIIPPLGLIYFFVQFPYVASSRQL	1140
DB	1081	DTVDSMIDPEVTKMPGSLFNFGVIGACIVILLATPAAIIIPPLGLIYFFVQFPYVASSRQL	1140
QY	1141	KRLSVSRSPVYSHFNELLGVSIVRAFEEOERFTHOSDLKVDENOKAYPSIVANRWLA	1200
DB	1141	KRLSVSRSPVYSHFNELLGVSIVRAFEEOERFTHOSDLKVDENOKAYPSIVANRWLA	1200
QY	1201	VRLECVGNCIVLFAALFAVISHRSLSAGLVGLSVSYSLQVTTYLNWLVRMSSEMETNIVA	1260
DB	1201	VRLECVGNCIVLFAALFAVISHRSLSAGLVGLSVSYSLQVTTYLNWLVRMSSEMETNIVA	1260
QY	1261	VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNYCLRYREDLDPVLRHINVTINGG	1320
DB	1261	VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNYCLRYREDLDPVLRHINVTINGG	1320
QY	1321	EKVGIVGRTGAGKSSLTGLFRINESAEGETIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380
DB	1321	EKVGIVGRTGAGKSSLTGLFRINESAEGETIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380
QY	1381	SGSLRMNLDPFSQYSDBEVWTSLELAHLKDFVSALPKLDHECAEGGENLSVGORQLVCL	1440
DB	1381	SGSLRMNLDPFSQYSDBEVWTSLELAHLKDFVSALPKLDHECAEGGENLSVGORQLVCL	1440
QY	1441	ARALLRTKILVLDEATAAVDLETTDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL	1500
DB	1441	ARALLRTKILVLDEATAAVDLETTDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL	1500
QY	1501	DKGIEQYEGAPSDLLQOQGLFYSNKADAGLV	1531
DB	1501	DKGIEQYEGAPSDLLQOQGLFYSNKADAGLV	1531

RESULT 9

ADB87954

ID ADB87954 standard; protein; 1531 AA.

AC ADB87954;

XX 04-DEC-2003 (first entry)

XX Human UGT1A1 protein sequence SEQ ID NO:678.

DE irinotecan; cancer; UGT1A1; cytostatic; topoisomerase I inhibitor;  
 XX colorectal cancer; cervical cancer; gastric cancer; lung cancer;  
 KW ovarian cancer; pancreatic cancer; malignant glioma;  
 KW uridine diphosphate glycosyltransferase I member A1.  
 XX Homo sapiens.

XX WO2003013536-A2.

FN 20-FEB-2003.

PD 23-JUL-2002; 2002WO-EP008217.

PP 23-JUL-2001; 2001EP-00117608.

XX 24-MAY-2002; 2002EP-00011710.

PR (EPID-) EPIDAUS BIOTECHNOLOGIE AG.

XX Heinrich G, Kerb R;

XX WPI; 2003-289896/28.



XX Use of irinotecan to treat cancer patient by determining if patient has  
PT variant alleles of UGT1A1 gene, administering increased/decreased amounts  
PT of irinotecan based on increased/decreased levels of UGT1A1 gene product.  
XX  
PS Disclosure; SEQ ID NO 678; 107pp; English.  
XX  
CC The invention relates to the novel use of irinotecan to treat a patient  
CC suffering from cancer. This involves determining if the patient has one  
CC or more variant alleles of the UGT1A1 gene, and if the patient has one  
CC more of such variant alleles, irinotecan is administered in an increased  
CC or decreased amount in comparison to the amount that is administered  
CC without regard to the patient's alleles in the UGT1A1 gene. The invention  
CC has cytostatic activity. A composition of the invention acts as a  
CC topoisomerase I inhibitor. The method is useful for treating a patient,  
CC an animal e.g. mouse or a human, preferably African or Asian, suffering  
CC from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,  
CC pancreatic cancer or malignant glioma. The present sequence is used in  
CC the exemplification of the invention.  
XX  
SQ Sequence 1531 AA;  
Query Match 79.4%; Score 7860; DB 7; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MALRGFCAGDSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFVLMACFPFPFLYLSRH 60  
DB 1 MALRGFCAGDSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFVLMACFPFPFLYLSRH 60  
QY 61 DRGYIQMTPLNKTALGFLWIVCWADLFYFWSRSGIFLAPVFLVSPDLLGTTLLA 120  
DB 61 DRGYIQMTPLNKTALGFLWIVCWADLFYFWSRSGIFLAPVFLVSPDLLGTTLLA 120  
QY 121 TFLIQERRKGVSQSGIMLTFHVLVLCALILRSKIMTALKEDAQVDLFRDITVYVFS 180  
DB 121 TFLIQERRKGVSQSGIMLTFHVLVLCALILRSKIMTALKEDAQVDLFRDITVYVFS 180  
QY 181 LLLIQLVLSGFCSDRPLSETIHDNPPCESSASFLSRTTFWITGLVGRVQPLEGSD 240  
DB 181 LLLIQLVLSGFCSDRPLSETIHDNPPCESSASFLSRTTFWITGLVGRVQPLEGSD 240  
QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKSSKVDANEVEAL 300  
DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKSSKVDANEVEAL 300  
QY 301 IVKSPQKWNPSLFVLYKTFPGYFLMSFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
DB 301 IVKSPQKWNPSLFVLYKTFPGYFLMSFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
QY 361 WQGYFVTVLLFVTAQLVHLYOYFHI CFVSGMRKTA VIGAVYKALVITNSARKSTV 420  
DB 361 WQGYFVTVLLFVTAQLVHLYOYFHI CFVSGMRKTA VIGAVYKALVITNSARKSTV 420  
QY 421 GEIVNLSVDAQRFMDLATYINMWSAPLQVILALYLLWNLGSPSLAGVAVMVLMPVN 480  
DB 421 GEIVNLSVDAQRFMDLATYINMWSAPLQVILALYLLWNLGSPSLAGVAVMVLMPVN 480  
QY 481 AVAMKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKYAWELAPKDKVLAIRQEBELKVLK 540  
DB 481 AVAMKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKYAWELAPKDKVLAIRQEBELKVLK 540  
QY 541 KSAYLSAVGTFTWCTPPLVALCTFAVYVITDENNILDQAOTAFVSLAFNLRFPNLILP 600  
DB 541 KSAYLSAVGTFTWCTPPLVALCTFAVYVITDENNILDQAOTAFVSLAFNLRFPNLILP 600  
QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGSITVTRNATFTWASDPT 660  
DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGSITVTRNATFTWASDPT 660  
QY 661 LNTGITSIPGALVAVGVQCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQWIND 720  
DB 661 LNTGITSIPGALVAVGVQCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQWIND 720

QY 721 SLRENILFGCOLEBPYRVSIVQACALLPDLBELPSGDRTEIGEGKWNLSGGQKQVSLAR 780  
DB 721 SLRENILFGCOLEBPYRVSIVQACALLPDLBELPSGDRTEIGEGKWNLSGGQKQVSLAR 780  
QY 781 AVYNSADIYLLPDDPLSAVDHVGKHI FENVIGPKMKNKTRILVTHSMYLPQVDVIV 840  
DB 781 AVYNSADIYLLPDDPLSAVDHVGKHI FENVIGPKMKNKTRILVTHSMYLPQVDVIV 840  
QY 841 MSGGKISMGYSQELLARDGAFELRYASTEQDABENGVTGVSFGPKAKQWENGM 900  
DB 841 MSGGKISMGYSQELLARDGAFELRYASTEQDABENGVTGVSFGPKAKQWENGM 900  
QY 901 LVTDSAGLQORQLSSSSSYSGDISRHHNSTAEQKAEKBEETWKLMEADKAQTCQVKL 960  
DB 901 LVTDSAGLQORQLSSSSSYSGDISRHHNSTAEQKAEKBEETWKLMEADKAQTCQVKL 960  
QY 961 SVYWDYKAIGLFTISFLSIFLMCNHVSALASNYWLSLWTDPIVNGTQEHTKVLSVYG 1020  
DB 961 SVYWDYKAIGLFTISFLSIFLMCNHVSALASNYWLSLWTDPIVNGTQEHTKVLSVYG 1020  
QY 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILRS PMSFFERTPSGNLVNRFSKEL 1080  
DB 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILRS PMSFFERTPSGNLVNRFSKEL 1080  
QY 1081 DTVDSMIPEVIKMFMSGLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
DB 1081 DTVDSMIPEVIKMFMSGLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
QY 1141 KRLSVSRSPVYSHNETLLGVSVIRAPBEOERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
DB 1141 KRLSVSRSPVYSHNETLLGVSVIRAPBEOERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
QY 1201 VRLECVGNCIVLPAALPAVISRHSISAGLVGLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
DB 1201 VRLECVGNCIVLPAALPAVISRHSISAGLVGLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
QY 1261 VERLKEYSETBEAPWQIQETAPPSSWPQVGRVFEPRNYCLARYREDLDVFLRHINVTINGG 1320  
DB 1261 VERLKEYSETBEAPWQIQETAPPSSWPQVGRVFEPRNYCLARYREDLDVFLRHINVTINGG 1320  
QY 1321 EKVGIIVGTGAGKSLTGLFRINESABGEIIDIINIAKIGLHDLRKFITIIIPQDPVLF 1380  
DB 1321 EKVGIIVGTGAGKSLTGLFRINESABGEIIDIINIAKIGLHDLRKFITIIIPQDPVLF 1380  
QY 1381 SGLSRMNLDPFSQYSDDEEVTSLSLAHUKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
DB 1381 SGLSRMNLDPFSQYSDDEEVTSLSLAHUKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
QY 1441 ARALLRKTILVLDDEATAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
DB 1441 ARALLRKTILVLDDEATAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
QY 1501 DKGEIQEYGAESDILLQORGLFYSMADAGLV 1531  
DB 1501 DKGEIQEYGAESDILLQORGLFYSMADAGLV 1531  
RESULT 10  
ADB96937  
ID ADB96937 standard; protein; 1531 AA.  
XX  
AC ADB96937;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Human MDR1 related protein sequence SEQ ID NO:678.  
XX  
KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;  
KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;  
KW multidrug resistance 1; MDR1; cytostatic; human; CYP3A5; MRP1; MDR1;  
XX TOPI.

OS Homo sapiens.  
 PN WO2003013537-A2.  
 XX 20-FEB-2003.  
 XX 23-JUL-2002; 2002WO-BF008219.  
 XX 23-JUL-2001; 2001EP-00117608.  
 PR 24-MAY-2002; 2002EP-00011710.  
 XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
 XX Heinrich G, Kerb R;  
 XX WPI; 2003-268145/26.  
 XX  
 FT New use of irinotecan for preparation of pharmaceutical compositions for  
 PT treating cancer in subject having genome with variant allele comprising  
 PT multidrug resistance 1 polynucleotide.  
 XX Disclosure; SEQ ID NO 678; 130pp; English.  
 XX  
 CC The invention relates to the novel use of irinotecan or its derivative  
 CC for the preparation of pharmaceutical compositions for treating  
 CC colorectal, gastric, lung, ovarian or pancreatic cancer, or  
 CC malignant glioma in a subject having a genome with a variant allele which  
 CC comprises a multidrug resistance 1 (MDR1) polynucleotide. A composition  
 CC of the invention has cytostatic activity. The invention is useful for the  
 CC preparation of pharmaceutical compositions for treating colorectal,  
 CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant  
 CC glioma in a subject (preferably human, more preferably African or Asian)  
 CC or a mouse. The present sequence is used in the exemplification of the  
 CC invention.  
 XX  
 SQ Sequence 1531 AA;  
 Query Match 79.4%; Score 7860; DB 7; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MALRFGCSADGSDPLDWNVTWNTSNPDTKCFQNTLVWVPCFVLMACFPFFFLYLSRH 60  
 1 MALRFGCSADGSDPLDWNVTWNTSNPDTKCFQNTLVWVPCFVLMACFPFFFLYLSRH 60  
 61 DRGYIQTPLNKTALGFLWVWADLFSWERSRGIFLAPVFLVSPITLGLITLLA 120  
 61 DRGYIQTPLNKTALGFLWVWADLFSWERSRGIFLAPVFLVSPITLGLITLLA 120  
 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 181 LLLIQLVLSCSFSDRSPLESETHDNPCESSASFLSRITFWITGLIVRGYRPLEGSD 240  
 181 LLLIQLVLSCSFSDRSPLESETHDNPCESSASFLSRITFWITGLIVRGYRPLEGSD 240  
 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPKVYVSSKDPAPKESKSDVANEVEAL 300  
 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPKVYVSSKDPAPKESKSDVANEVEAL 300  
 301 IVKSPQKEWNPSPFLKVLVYKTEGYPFLMSFFPKAIHDLMMFSGPQLKLLIKFVNDTKAPD 360  
 301 IVKSPQKEWNPSPFLKVLVYKTEGYPFLMSFFPKAIHDLMMFSGPQLKLLIKFVNDTKAPD 360  
 361 WQGYEYTVLLFVTAQLQTLVHLQVPHICFVSGMRKTAIVGAVRKALVITNSARKSSTV 420  
 361 WQGYEYTVLLFVTAQLQTLVHLQVPHICFVSGMRKTAIVGAVRKALVITNSARKSSTV 420  
 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLMLNLGFSVLGAVVWMLVMPVN 480  
 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLMLNLGFSVLGAVVWMLVMPVN 480

QY 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROELKVLK 540  
 DB 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROELKVLK 540  
 QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAQAFVLSALFNILRPFPLILP 600  
 DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAQAFVLSALFNILRPFPLILP 600  
 QY 601 MVISSIVQASVSLAKRLRIFLSHEELEPDSIERRPVKDGGGTSITVRNATTTWARSDDPT 660  
 DB 601 MVISSIVQASVSLAKRLRIFLSHEELEPDSIERRPVKDGGGTSITVRNATTTWARSDDPT 660  
 QY 661 LINGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAKGVSVAVYVPOQAWIQND 720  
 DB 661 LINGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAKGVSVAVYVPOQAWIQND 720  
 QY 721 SLRENILFGCOLEBPYRVSIVIOACALLPDLBILPSGDRTEIGEGVNLSSGQKQVSLAR 780  
 DB 721 SLRENILFGCOLEBPYRVSIVIOACALLPDLBILPSGDRTEIGEGVNLSSGQKQVSLAR 780  
 QY 781 AVYSNADIYLPDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840  
 DB 781 AVYSNADIYLPDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840  
 QY 841 MSGGKISEMSYQBELLDAGAFABFLRTYASTEQEADAEENGVTGVSQPGKEAKQMGCM 900  
 DB 841 MSGGKISEMSYQBELLDAGAFABFLRTYASTEQEADAEENGVTGVSQPGKEAKQMGCM 900  
 QY 901 LVTSAGKQLOQRLSSSSSYSGDISRHNNSTAEIQAFAKKEETWKLMEADKAQTQVKL 960  
 DB 901 LVTSAGKQLOQRLSSSSSYSGDISRHNNSTAEIQAFAKKEETWKLMEADKAQTQVKL 960  
 QY 961 SVYDYNKAIGLFTSLFSLFMCNHSALASNLWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 DB 961 SVYDYNKAIGLFTSLFSLFMCNHSALASNLWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 QY 1021 ALGISQIAGVFGSMAYSIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL 1080  
 DB 1021 ALGISQIAGVFGSMAYSIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKEL 1080  
 QY 1081 DTVDSPITEVTKPMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140  
 DB 1081 DTVDSPITEVTKPMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140  
 QY 1141 KRLESVSRSPYSHFNFTLLGVSVIRAFEBQERFIHOSDKVDENOKAYYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPYSHFNFTLLGVSVIRAFEBQERFIHOSDKVDENOKAYYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVLSYSYSLQVTTYLNWLVRMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVLSYSYSLQVTTYLNWLVRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKEAPWIOETAPPSSWPQVGRVEFRNYCLRYRDLDFVLRIHINVTNGG 1320  
 DB 1261 VERLKEYSETEKEAPWIOETAPPSSWPQVGRVEFRNYCLRYRDLDFVLRIHINVTNGG 1320  
 QY 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 DB 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGSLRMNLDPFQSYSDSEVWTSLELAHKDFVSALPKDLHDECAEGGENLSVGQRLVCL 1440  
 DB 1381 SGSLRMNLDPFQSYSDSEVWTSLELAHKDFVSALPKDLHDECAEGGENLSVGQRLVCL 1440  
 QY 1441 ARALLRKTILVLDDEATAVNDLETDLLIQTIRTOFEDCTVLTIAHRLNIMDYTRIVL 1500  
 DB 1441 ARALLRKTILVLDDEATAVNDLETDLLIQTIRTOFEDCTVLTIAHRLNIMDYTRIVL 1500  
 QY 1501 DKGEIQBYGAPSDLLQORGLFYSMKADAGLV 1531  
 DB 1501 DKGEIQBYGAPSDLLQORGLFYSMKADAGLV 1531

RESULT 11	Db	301	IVKSPQKEMNPSLFKVLKYKTFGPYFLMSFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD	360
AD92128	Qy	361	WQGYFTVLLFVTACLOTLVHLQYFHICFVSGMIRKTAIVIGAVYRKALVITNSARKSTV	420
ID ADB92128 standard; protein; 1531 AA.	Db	361	WQGYFTVLLFVTACLOTLVHLQYFHICFVSGMIRKTAIVIGAVYRKALVITNSARKSTV	420
AC ADB92128;	Qy	421	GEIYNLMSVDAQRPMDLATYINMIWSAPLQVILALYLLNLNLGSPSVLAGVAVMLMVPV	480
04-DEC-2003 (first entry)	Db	421	GEIYNLMSVDAQRPMDLATYINMIWSAPLQVILALYLLNLNLGSPSVLAGVAVMLMVPV	480
Human MDR1 related protein sequence SEQ ID NO:678.	Qy	481	AVMAKTKTYQVAHMKSKDNRIKLMNEITLNGIKVLKLYAWELAFKDKVLAIRQEBELKVLK	540
irinotecan; colorectal cancer; cervical cancer; gastric cancer;	Db	481	AVMAKTKTYQVAHMKSKDNRIKLMNEITLNGIKVLKLYAWELAFKDKVLAIRQEBELKVLK	540
lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;	Qy	541	KSAYLSAVGTFTWCTPFLVALCTFAVVYVITDENNILDQAQAFVLSLALFNILRFFLNILP	600
multidrug resistance 1; MDR1; cytostatic; human; UGT1A1; MRP1; TOP1.	Db	541	KSAYLSAVGTFTWCTPFLVALCTFAVVYVITDENNILDQAQAFVLSLALFNILRFFLNILP	600
Homo sapiens.	Qy	601	MVISIIVQASVSLKRLRIFLSHEBLEPDSIERRPVKGGGTSITVRNATFTWARSDDPT	660
WO2003013535-A2.	Db	601	MVISIIVQASVSLKRLRIFLSHEBLEPDSIERRPVKGGGTSITVRNATFTWARSDDPT	660
20-FEB-2003.	Qy	661	LNGITFSIPEGALVAVVGVCGCKSSLLSALLAEMDKVEGHVAKGVSAYVVPQQAIIQND	720
23-JUL-2002; 2002WO-EP008220.	Db	661	LNGITFSIPEGALVAVVGVCGCKSSLLSALLAEMDKVEGHVAKGVSAYVVPQQAIIQND	720
23-JUL-2001; 2001EP-00117608.	Qy	721	SLRENILFGCOLERPYRSVIOACALLPDLEILPSGDRTEIGEKNVLSGGQKQVSLAR	780
24-MAY-2002; 2002EP-00011710.	Db	721	SLRENILFGCOLERPYRSVIOACALLPDLEILPSGDRTEIGEKNVLSGGQKQVSLAR	780
(EPID-) EPIDAUROS BIOTECHNOLOGIE AG.	Qy	781	AVYGNADIYLPDDPLSAVDHVGKHI PENVI GPKGMLKNKTRILVTHSMVSLPQVDVIIV	840
Heinrich G, Korb R;	Db	781	AVYGNADIYLPDDPLSAVDHVGKHI PENVI GPKGMLKNKTRILVTHSMVSLPQVDVIIV	840
WPI; 2003-342400/32.	Qy	841	MSGKISMSGYSQELLARDGAFBFLRTYASTEQDAEENGVTGSGPGKEAKOMENGM	900
New use of irinotecan for preparation of pharmaceutical compositions for	Db	841	MSGKISMSGYSQELLARDGAFBFLRTYASTEQDAEENGVTGSGPGKEAKOMENGM	900
treating cancer in subject having genome with variant allele comprising	Qy	901	LVTDSAGKQLQRLSSSSYSYSGDISRHNSTAEQKAEKKEETWKLMEADKAQIQVKL	960
multidrug resistance 1 polynucleotide.	Db	901	LVTDSAGKQLQRLSSSSYSYSGDISRHNSTAEQKAEKKEETWKLMEADKAQIQVKL	960
Disclosure; SEQ ID NO 678; 104pp; English.	Qy	961	SVYDYMKAIGLFTSIFLSIFLFCMCHVSALASNYWLSLWTDPIVNGTQHTKVLRSVYG	1020
The invention relates to a novel use of irinotecan or its derivative for	Db	961	SVYDYMKAIGLFTSIFLSIFLFCMCHVSALASNYWLSLWTDPIVNGTQHTKVLRSVYG	1020
the preparation of a pharmaceutical composition for treating colorectal,	Qy	1021	ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTSGNLVNRFSKEL	1080
cervical, gastric, lung, ovarian or pancreatic cancer, or malignant	Db	1021	ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTSGNLVNRFSKEL	1080
glioma in a subject having a genome with a variant allele which comprises	Qy	1081	DTVDSMIPEVTKMFMSGLFNIVIGACIVILLATPIAAIIIPGLGLYFFQRFYVASSROL	1140
a multidrug resistance 1 (MDR1) polynucleotide. A composition of the	Db	1081	DTVDSMIPEVTKMFMSGLFNIVIGACIVILLATPIAAIIIPGLGLYFFQRFYVASSROL	1140
invention has cytostatic activity. The present sequence is used in the	Qy	1141	KRLSESVRSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA	1200
exemplification of the invention.	Db	1141	KRLSESVRSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA	1200
Seq Sequence 1531 AA;	Qy	1201	VRLECVGNCIVLFAALFAVIRSRHSLSAGLVGSYSLOVTTYNLWLRMSSEMETNIVA	1260
Query Match 79.4%; Score 7860; DB 7; Length 1531;	Db	1201	VRLECVGNCIVLFAALFAVIRSRHSLSAGLVGSYSLOVTTYNLWLRMSSEMETNIVA	1260
Best Local Similarity 100.0%; Pred. No. 0;	Qy	1261	VERLKEYSETKEAPWQIQETAPSSWPQVGRVEFRNYCLARYREDLDFVLRHINTVINGG	1320
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Db	1261	VERLKEYSETKEAPWQIQETAPSSWPQVGRVEFRNYCLARYREDLDFVLRHINTVINGG	1320
1 MALRGFCSDGSDPLWDNVTWNTSNPDTKCFQNTVLVWVPCFYLWACFPFFFLYLSRH	Qy	1321	EKVGVGRTGAGKSLTLGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPDPVLF	1380
1 MALRGFCSDGSDPLWDNVTWNTSNPDTKCFQNTVLVWVPCFYLWACFPFFFLYLSRH	Db	1321	EKVGVGRTGAGKSLTLGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPDPVLF	1380
61 DRGYIOMPLNKTALGFLMWVCWADLFYSFWSRSGIFLAPVFLVSPILLGTTLLA	Qy	1381	SGSLRMNLDPPSOVSDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL	1440
61 DRGYIOMPLNKTALGFLMWVCWADLFYSFWSRSGIFLAPVFLVSPILLGTTLLA	Db	1381	SGSLRMNLDPPSOVSDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL	1440
121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS	Qy			
121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS	Db			
181 LLLIQLVLCFSDRSPFLSETHDNPCESSASFLSRITFWITGLIVRGVRLQEGSD	Qy			
181 LLLIQLVLCFSDRSPFLSETHDNPCESSASFLSRITFWITGLIVRGVRLQEGSD	Db			
241 LWSLNKEDTSEOVVPLVKNWKECAKTRKQPVVYSSKDPQPKSSKVDANEVEAL	Qy			
241 LWSLNKEDTSEOVVPLVKNWKECAKTRKQPVVYSSKDPQPKSSKVDANEVEAL	Db			
301 IVKSPQKEMNPSLFKVLKYKTFGPYFLMSFFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD	Qy			

QY 1441 ARALLRKTILVDEATAVDLETDLIQSTIRTOPECTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTILVDEATAVDLETDLIQSTIRTOPECTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEYGAPSDLLQORGLFYFMAKDAGLV 1531  
 Db 1501 DKGEIQEYGAPSDLLQORGLFYFMAKDAGLV 1531

RESULT 12  
 ID ADD44764  
 XX ADD44764 standard; protein; 1531 AA.  
 AC ADD44764;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human Protein P33527, SEQ ID NO 10193.  
 XX  
 KW Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 14-AUG-2002; 2002WO-US025765.  
 XX  
 PR 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-034382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX  
 PA (GEO ) GEN HOSPITAL CORP.  
 PA (FARB ) BAYER AG.  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX  
 DR WPI; 2003-268312/26.  
 DR GENBANK; P33527.  
 XX  
 PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 XX  
 PS Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in table 2 of the specification) which is differentially expressed during pain. Note:

CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 1531 AA;  
 Query Match 79.4%; Score 7860; DB 7; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGFCSADGSDPLMDWNTWNTSNPDTKCFQNTVLVWVPCFYLWACFFFYLYLSRH 60  
 Db 1 MALRGFCSADGSDPLMDWNTWNTSNPDTKCFQNTVLVWVPCFYLWACFFFYLYLSRH 60  
 QY 61 DRGYIOMTPLNKTALGFLFWIVCWADLFYFWSRSGIFLAPVFLVSPLLGLTLLA 120  
 Db 61 DRGYIOMTPLNKTALGFLFWIVCWADLFYFWSRSGIFLAPVFLVSPLLGLTLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLPRDITFYVYFS 180  
 Db 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLPRDITFYVYFS 180  
 QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPSESSASFLSRITFWITGLIVRGYRPLEGSD 240  
 Db 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPSESSASFLSRITFWITGLIVRGYRPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKKCAKTRQPKVVTYSSKDPAPQKSSSKVDANEVEAL 300  
 Db 241 LWSLNKEDTSEQVVPVLVKNWKKCAKTRQPKVVTYSSKDPAPQKSSSKVDANEVEAL 300  
 QY 301 IVKSPQKEWNPSPKLVYKTFPGPYFLMSFPFKAIDHLMFSGPOILKLIKFNVDNTRAPD 360  
 Db 301 IVKSPQKEWNPSPKLVYKTFPGPYFLMSFPFKAIDHLMFSGPOILKLIKFNVDNTRAPD 360  
 QY 361 WQGYFTVLLFVTACLOTLVLHQYFHCIFVSGMRKITAIVGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFTVLLFVTACLOTLVLHQYFHCIFVSGMRKITAIVGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLSVDAOREFMDLATYINMIWSAPLOVILALYLLNLNLSGLAGVAVWVLMVFN 480  
 Db 421 GEIVNLSVDAOREFMDLATYINMIWSAPLOVILALYLLNLNLSGLAGVAVWVLMVFN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTFTWVCTPFLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRPLNLP 600  
 Db 541 KSAYLSAVGTFTWVCTPFLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRPLNLP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGCGTNSITVSNATFTWARSPPPT 660  
 Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGCGTNSITVSNATFTWARSPPPT 660  
 QY 661 LNGTHIPGALVAVVGVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWTQND 720  
 Db 661 LNGTHIPGALVAVVGVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWTQND 720  
 QY 721 SLRENILFGCLBEPYRVSIVIQACALLPDLEILPSGDRTEIGEGVNLSSGQKQVSLAR 780  
 Db 721 SLRENILFGCLBEPYRVSIVIQACALLPDLEILPSGDRTEIGEGVNLSSGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIKPKMLKNNKTRILVTHSMYSVLPQVDVIV 840  
 Db 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIKPKMLKNNKTRILVTHSMYSVLPQVDVIV 840  
 QY 841 MSGKISEMGSYQELLARDGAFELFTYASTEQEQDAEENGVTGVSQPGKEAKQEMNGM 900  
 Db 841 MSGKISEMGSYQELLARDGAFELFTYASTEQEQDAEENGVTGVSQPGKEAKQEMNGM 900  
 QY 901 LVTDGAKQLQRLSSSSSSSGDISRHNSHTAELOKAEKKEETWKLMEADKAQTQVKL 960  
 Db 901 LVTDGAKQLQRLSSSSSSSGDISRHNSHTAELOKAEKKEETWKLMEADKAQTQVKL 960

QY 961 SVYDYMKAIGLFTISFLSIFLPMCHVNSALASNYWLSLWTDPIVNGTOEHTKVLRSYVG 1020  
DB 961 SVYDYMKAIGLFTISFLSIFLPMCHVNSALASNYWLSLWTDPIVNGTOEHTKVLRSYVG 1020  
QY 1021 ALGISQITAVGYSMAVSGIGGLASRCLHVDLLHSIIRSPMSFFERTSGNLVNRFSKEL 1080  
DB 1021 ALGISQITAVGYSMAVSGIGGLASRCLHVDLLHSIIRSPMSFFERTSGNLVNRFSKEL 1080  
QY 1081 DTVDSMIEVIMKMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
DB 1081 DTVDSMIEVIMKMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
QY 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
DB 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
QY 1201 VRLECVGNICIVLFAALFAVISHSLSAGLVGSYSISQVTTYNLWLRMSSEMETNIVA 1260  
DB 1201 VRLECVGNICIVLFAALFAVISHSLSAGLVGSYSISQVTTYNLWLRMSSEMETNIVA 1260  
QY 1261 VERLKEYSETEKAPQIETAPPSSWQVGRVFNRYCLRYREDLDFVLRHINVTINGG 1320  
DB 1261 VERLKEYSETEKAPQIETAPPSSWQVGRVFNRYCLRYREDLDFVLRHINVTINGG 1320  
QY 1321 EKVIGVGRGTAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
DB 1321 EKVIGVGRGTAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
QY 1381 SGLSRMNLDPFSQYDEEVTWTSLELAHLKDFVSALPKLDHECARGGENLSVGQRQLVCL 1440  
DB 1381 SGLSRMNLDPFSQYDEEVTWTSLELAHLKDFVSALPKLDHECARGGENLSVGQRQLVCL 1440  
QY 1441 ARALARKTILVLDATAAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
DB 1441 ARALARKTILVLDATAAVDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
QY 1501 DKGEIQEYGAPELDLQORGLFYSMADAGLV 1531  
DB 1501 DKGEIQEYGAPELDLQORGLFYSMADAGLV 1531

RESULT 13

ADN97111  
ID ADN97111 standard; protein; 1531 AA.  
XX AC ADN97111;  
XX DT 01-JUL-2004 (first entry)  
XX DE MRP1 protein.  
XX KW multiple drug resistance protein; MRP; Drosophila melanogaster;  
XX KW Anopheles gambiae; insecticide.  
XX OS Homo sapiens.  
XX FN WO2004029088-A2.  
XX PD 08-APR-2004.  
XX PF 25-SEP-2003; 2003WO-EP012400.  
XX PR 26-SEP-2002; 2002US-0413469P.  
XX XX (INSP ) INST PASTEUR.  
XX PA (CNRS ) CENT NAT RECH SCI.  
XX PI Roth CW, Brey PT, Holm I, Graillies M, Rzhetsky A;  
XX WPI; 2004-305150/28.  
XX NT New polynucleotide sequence encoding multiple drug resistance proteins

PT from Drosophila melanogaster or Anopheles gambiae, useful in developing  
XX effective insecticides.  
PS Claim 1; SEQ ID NO 6; 58pp; English.  
XX The present invention relates to a purified polynucleotide or its  
CC fragment and comprises a sequence encoding multiple drug resistance  
CC proteins (MRPs) from Drosophila melanogaster or Anopheles gambiae. The  
CC polynucleotide is useful in developing effective insecticides. The  
CC present sequence represents human MRP1.  
XX SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 8; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MALRGFCSADGSDPLMDWNTWNTSNPDFTKCFONTVLVWVPCFVLMACFPFYLGRH 60  
DB 1 MALRGFCSADGSDPLMDWNTWNTSNPDFTKCFONTVLVWVPCFVLMACFPFYLGRH 60  
QY 61 DRGIQMTPLNKTALGFLIWCWADLFYSFWSRSGIFLAPVFLVSPILLGTTLLA 120  
DB 61 DRGIQMTPLNKTALGFLIWCWADLFYSFWSRSGIFLAPVFLVSPILLGTTLLA 120  
QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVFS 180  
DB 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVFS 180  
QY 181 LLLQLVLSCFSDRSPLFSETHDNPSPSSASFLSRITFWITGLIVRGVRLQEGSD 240  
DB 181 LLLQLVLSCFSDRSPLFSETHDNPSPSSASFLSRITFWITGLIVRGVRLQEGSD 240  
QY 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKOPKVYVSSKDPAPQPKESKVDANEVEAL 300  
DB 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKOPKVYVSSKDPAPQPKESKVDANEVEAL 300  
QY 301 IVKSPQKEWNPFLFKVLYKTFPGPYFLMSFFFKAIHDLMMFSGPQILKLIKFVNDTKAPD 360  
DB 301 IVKSPQKEWNPFLFKVLYKTFPGPYFLMSFFFKAIHDLMMFSGPQILKLIKFVNDTKAPD 360  
QY 361 WQGYFYTVLLFWTACQLVLHVOYFHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
DB 361 WQGYFYTVLLFWTACQLVLHVOYFHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLWNLGPSVLGAVAVMLVMPVN 480  
DB 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLWNLGPSVLGAVAVMLVMPVN 480  
QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROEELKVLK 540  
DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROEELKVLK 540  
QY 541 KSAVLSAVGTTTWTCTPPLVALCTPAVVTVDENNILDAQTAFVSLALFNILRFPNLILP 600  
DB 541 KSAVLSAVGTTTWTCTPPLVALCTPAVVTVDENNILDAQTAFVSLALFNILRFPNLILP 600  
QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660  
DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660  
QY 661 LINGITFISPEGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720  
DB 661 LINGITFISPEGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720  
QY 721 SLRENILFGCOLEEPYRSVIOACALLPDLPEILPSGDRTEIGEKGVNLSGGQKQVSLAR 780  
DB 721 SLRENILFGCOLEEPYRSVIOACALLPDLPEILPSGDRTEIGEKGVNLSGGQKQVSLAR 780  
QY 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIGPKMLKNTKILVTHSNLYLPQVDVILV 840  
DB 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIGPKMLKNTKILVTHSNLYLPQVDVILV 840

```
QY 841 MSGGKISEMGYSQELLARDGAFABELRTYASTEQDQDAEENGVTGVSQPGKEAKOMENG 900
Db 841 MSGGKISEMGYSQELLARDGAFABELRTYASTEQDQDAEENGVTGVSQPGKEAKOMENG 900
QY 901 LVTSAGKQORQLSSSSSYSGDISRHNSHNSAEQKAEKKEETWKLMEADKAQTQGVKL 960
Db 901 LVTSAGKQORQLSSSSSYSGDISRHNSHNSAEQKAEKKEETWKLMEADKAQTQGVKL 960
QY 961 SVYDYMKAIGLFTSFLIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKVRLSVYG 1020
Db 961 SVYDYMKAIGLFTSFLIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKVRLSVYG 1020
QY 1021 ALGISQIAGVGYMAVIGIGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKEL 1080
Db 1021 ALGISQIAGVGYMAVIGIGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKEL 1080
QY 1081 DTVDMSIPEVIMFMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQFYVASSRQL 1140
Db 1081 DTVDMSIPEVIMFMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQFYVASSRQL 1140
QY 1141 KRLESVSRSVPYSHENETLLGVSVIRAFERQERFTHQSDLVKVDENQKAYYSIVANRWLA 1200
Db 1141 KRLESVSRSVPYSHENETLLGVSVIRAFERQERFTHQSDLVKVDENQKAYYSIVANRWLA 1200
QY 1201 VRLECVGNCIVLFAALFAVIRSHLSAGLVLSYSLSQVTTYLNWLVRMSSEMETNIVA 1260
Db 1201 VRLECVGNCIVLFAALFAVIRSHLSAGLVLSYSLSQVTTYLNWLVRMSSEMETNIVA 1260
QY 1261 VERLUKEYETKEAPWQIQTAPPSSWPQGRVFRNVCYLYRBDLPVLRHINVTINGG 1320
Db 1261 VERLUKEYETKEAPWQIQTAPPSSWPQGRVFRNVCYLYRBDLPVLRHINVTINGG 1320
QY 1321 EKVGVIGRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRDKITIIIPQDPVLF 1380
Db 1321 EKVGVIGRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRDKITIIIPQDPVLF 1380
QY 1381 SGSLRMNLDPFSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAEGGNLSVGQRQLVCL 1440
Db 1381 SGSLRMNLDPFSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAEGGNLSVGQRQLVCL 1440
QY 1441 ARALLRKTKILVLDEATAVLETDLLIQSTIRTOQFEDCTVLTIAHRLNTIMDYTRVIL 1500
Db 1441 ARALLRKTKILVLDEATAVLETDLLIQSTIRTOQFEDCTVLTIAHRLNTIMDYTRVIL 1500
QY 1501 DKGEIQEYGAQSDLLQQRGLFYSMAKDAGLV 1531
Db 1501 DKGEIQEYGAQSDLLQQRGLFYSMAKDAGLV 1531

RESULT 14
ADN95929
ID ADN95929 standard; protein; 1530 AA.
XX AC ADN95929;
XX DT
XX DE Human BEC/LEC-related protein sequence SeqID853.
XX
KW growth; differentiation; blood endothelial cell; BEC;
KW lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;
KW lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;
KW vasotrophic; antiinflammatory; gene therapy; endothelial cell disorder;
KW inflammatory disease; cancer metastasis; lymphatic system; human.
XX
OS Homo sapiens.
XX
FN WO2003080640-A1.
XX
PD 02-OCT-2003.
XX
PF 07-MAR-2003; 2003WO-US0006900.
XX
```

```
PR 07-MAR-2002; 2002US-0363019P.
XX (LUDM-) LUDWIG INST CANCER RES.
PA (LICN) LICENTIA LTD.
XX
XX Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;
XX WPI: 2003-876899/81.
XX N-PSDB; ADN95930.
XX
XX Example 1; SEQ ID NO 853; 176pp; English.
XX
XX This invention relates to a method of differentially modulating the
XX growth or differentiation of blood endothelial cells (BEC) or lymphatic
XX endothelial cells (LEC) comprises contacting endothelial cells with a
XX composition comprising an agent that differentially modulates blood or
XX lymphatic endothelial cells. Treating hereditary lymphoedema comprises
XX identifying a human subject with lymphoedema and with a mutation in at
XX least one allele of a gene encoding a LEC protein, where the mutation
XX correlates with lymphoedema in human subjects, and with the proviso that
XX the LEC protein is not VEGFR-3; and administering to the subject a
XX composition comprising a lymphatic growth agent selected from VEGF-C or
XX VEGF-D polypeptides and polynucleotides. The invention may be useful for
XX the development of compounds with an antiangiogenic, cytostatic,
XX vasotrophic or antiinflammatory activity or for gene therapy. The method
XX is useful in modulating the growth or differentiation of blood
XX endothelial cells or lymphatic endothelial cells, in treating hereditary
XX lymphoedema, in screening for an endothelial cell disorder or
XX predisposition to the disorder or in monitoring the efficacy or toxicity
XX of a drug on endothelial cells. The agent is useful in manufacturing a
XX medicament for the differential modulation of blood vessel endothelial
XX cell or lymphatic vessel endothelial cell growth or differentiation. The
XX lymphatic growth agent may also be used in manufacturing a medicament for
XX the treatment of hereditary lymphoedema resulting from a mutation in a
XX LEC gene or of other diseases involving the lymphatic vessels, such as
XX various inflammatory diseases and cancer metastasis via the lymphatic
XX system. The present sequence is that of a human LEC/BEC differentially
XX expressed protein which is related to the method of the invention. Note:
XX This sequence does not appear in the specification but was obtained by
XX the indexer using the source data given in table 14 of the specification.
XX
XX Sequence 1530 AA;
XX
XX Query Match 79.3%; Score 7856; DB 7; Length 1530;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLMDWNVTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFFLYLSRH 60
Db 1 MALRGFCSADGSDPLMDWNVTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFFLYLSRH 60
QY 61 DRGYIOMTPLNKTALGFLLIWVCWADLVSFWSERSRGIFLAPVFLVSPFLIGITTLA 120
Db 61 DRGYIOMTPLNKTALGFLLIWVCWADLVSFWSERSRGIFLAPVFLVSPFLIGITTLA 120
QY 121 TFLIQLERRKGVSQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLPFDITFYVYFS 180
Db 121 TFLIQLERRKGVSQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLPFDITFYVYFS 180
QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPSSASFLSRITFWWTITGLIVRGYRPLEGSD 240
Db 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPSSASFLSRITFWWTITGLIVRGYRPLEGSD 240
QY 241 LWSLNKEDTSEQVVPVLVKWKKKCAKTRQPKVYVSSKDPKQPKSSKVDANEEVEAL 300
Db 241 LWSLNKEDTSEQVVPVLVKWKKKCAKTRQPKVYVSSKDPKQPKSSKVDANEEVEAL 300
QY 301 IVKSPQKEWNPFLFKVLYKTFPGYFLMSFFFKALHDLMMFSGPOLIKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEWNPFLFKVLYKTFPGYFLMSFFFKALHDLMMFSGPOLIKLLIKFVNDTKAPD 360
QY 361 WQGYFTYVLLFVTTACLOTVLVHQYFHCIFVSGMRKITAVTGANVRKALVITNSARKSTV 420
Db 361 WQGYFTYVLLFVTTACLOTVLVHQYFHCIFVSGMRKITAVTGANVRKALVITNSARKSTV 420
```

Db 361 WGYEYTVLLFVTAQLTLVLHQLYFHCIFVSGMRKTAIVAGVYKALVITNSARKSSTV 420  
QY 421 GEIVNLSVDAQFMDLATYINMIWSAPLQVILALYLWLNLPSPVLAGVAVMLVMPVN 480  
Db 421 GEIVNLSVDAQFMDLATYINMIWSAPLQVILALYLWLNLPSPVLAGVAVMLVMPVN 480  
QY 481 AVAMKTKTYQVAHKKSDNRKLMNEILNGIKVLKYAWELAPKDKVLAIROBELKVLK 540  
Db 481 AVAMKTKTYQVAHKKSDNRKLMNEILNGIKVLKYAWELAPKDKVLAIROBELKVLK 540  
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQCTAFVLSLALFNILRPPNLILP 600  
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQCTAFVLSLALFNILRPPNLILP 600  
QY 601 MVISSIVQASVSLKRLRIFLSHEELEPPSIERRPVKDGCGTNSITVRNATFTWARSDDPT 660  
Db 601 MVISSIVQASVSLKRLRIFLSHEELEPPSIERRPVKDGCGTNSITVRNATFTWARSDDPT 660  
QY 661 LNGITFPIPEGALVAVGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
Db 661 LNGITFPIPEGALVAVGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
QY 721 SIRENILFGCQLEBPYYSVIOACALLPDLEILPSGDRTEIGKGNVLSGGQKQVSLAR 780  
Db 721 SIRENILFGCQLEBPYYSVIOACALLPDLEILPSGDRTEIGKGNVLSGGQKQVSLAR 780  
QY 781 AVYSNADIYLFDDPLSADAVHGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840  
Db 781 AVYSNADIYLFDDPLSADAVHGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840  
QY 841 MSGGKISEMGYSQELLARDGAPAEFLRTYASTEQDABENGVTGVSFGKGAQKQNGM 900  
Db 841 MSGGKISEMGYSQELLARDGAPAEFLRTYASTEQDABENGVTGVSFGKGAQKQNGM 900  
QY 901 LVTDGAGLQRLQSSSSSYSGDISRHNSHNSAELOKAEAKKEETWKLMEADKAQGTQVKL 960  
Db 901 LVTDGAGLQRLQSSSSSYSGDISRHNSHNSAELOKAEAKKEETWKLMEADKAQGTQVKL 960  
QY 961 SVYDYMKAIGLFIISFLIFPMCNHVSALASNYWLSLWTDPIVNGTQHTKVLRSVYG 1020  
Db 961 SVYDYMKAIGLFIISFLIFPMCNHVSALASNYWLSLWTDPIVNGTQHTKVLRSVYG 1020  
QY 1021 ALGISQGIAGVGYSMVAISGIGLASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKEL 1080  
Db 1021 ALGISQGIAGVGYSMVAISGIGLASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKEL 1080  
QY 1081 DTVDMSIPEVIMKFMGSLFNIVIGACIVILLATPIAAIIIPGLGLIYFFVQFPYVASSRQL 1140  
Db 1081 DTVDMSIPEVIMKFMGSLFNIVIGACIVILLATPIAAIIIPGLGLIYFFVQFPYVASSRQL 1140  
QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFTHQSDLKVDENOKAYPSIVANRWLA 1200  
Db 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFTHQSDLKVDENOKAYPSIVANRWLA 1200  
QY 1201 VRLECVGNCIVLFAALFAVIRSHSLSAGLVGLSVSYSLQVTTYLNLWVRMSSEMETNIVA 1260  
Db 1201 VRLECVGNCIVLFAALFAVIRSHSLSAGLVGLSVSYSLQVTTYLNLWVRMSSEMETNIVA 1260  
QY 1261 VERLKEYSETKEAPWQIQTAPPSPVQVGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320  
Db 1261 VERLKEYSETKEAPWQIQTAPPSPVQVGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320  
QY 1321 EKVGVIGRTGAGKSSLTGLFRINESAGEIIIDGINIAKILGHDLRFRKTIIPQDPVLF 1380  
Db 1321 EKVGVIGRTGAGKSSLTGLFRINESAGEIIIDGINIAKILGHDLRFRKTIIPQDPVLF 1380  
QY 1381 SGSLRMNLDPFSQYSDVEEWTSLAHLKDFVSALPDKLDHECAGGENLSVGQRQLVCL 1440  
Db 1381 SGSLRMNLDPFSQYSDVEEWTSLAHLKDFVSALPDKLDHECAGGENLSVGQRQLVCL 1440  
QY 1441 ARALLRKTILVLDATAVLDLTDLLTQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Db 1441 ARALLRKTILVLDATAVLDLTDLLTQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500

QY 1501 DKGEIQEYGAPESDLLQQRGLFYSMKADAGL 1530  
Db 1501 DKGEIQEYGAPESDLLQQRGLFYSMKADAGL 1530  
RESULT 15  
AARS4928  
ID AARS4928 standard; protein; 1531 AA.  
XX AC AARS4928;  
XX DT 25-MAR-2003 (revised)  
XX DT 14-OCT-1994 (first entry)  
XX DE Multidrug resistance protein.  
XX KW Multidrug resistance protein; MRP; H69AR; cancer cell line; stem cell;  
XX KW cardiac muscle; transgenic animal.  
XX OS Homo sapiens.  
XX PN WO9410303-A1.  
XX PD 11-MAY-1994.  
XX PF 27-OCT-1993; 93WO-CA000439.  
XX PR 27-OCT-1992; 92US-00966923.  
XX PR 08-MAR-1993; 93US-00029340.  
XX PA (TOOH ) UNIV QUEBENS KINGSTON.  
XX PI Deesley RG, Cole SPC;  
XX DR WPI; 1994-167460/20.  
XX DR N-PSDB; AAQ65377.  
XX PT Multi-drug resistance gene - encodes protein capable of conferring multi-  
XX PT drug resistance on cells, useful in diagnostic and treatment methods.  
XX PS Disclosure; Page 69-74; 101pp; English.  
XX CC The multidrug resistant cancer cell line H69AR (ATCC CRL 11350) was used  
XX CC to identify cDNA encoding a novel protein associated with multidrug  
XX CC resistance, MRP. MRP may be expressed in e.g. hematopoietic stem cells or  
XX CC cardiac muscle, or in transgenic animals, or can be used to raise  
XX CC antibodies. (Updated on 25-MAR-2003 to correct PN field.)  
XX SQ Sequence 1531 AA;  
Query Match 79.3%; Score 7849; DB 2; Length 1531;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MALRGFCADGSDPLDNDWNTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60  
Db 1 MALRGFCADGSDPLDNDWNTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60  
QY 61 DRGYIQTMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPDLLGTTLLA 120  
Db 61 DRGYIQTMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPDLLGTTLLA 120  
QY 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILSKINTALKEDAQVDLFRDITFYVYFS 180  
Db 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILSKINTALKEDAQVDLFRDITFYVYFS 180  
QY 181 LLLIQLVLSCFSDRSPLFSETIHDPNPCPSSASFLSRITFWITGLIVRGVROPLEGSD 240  
Db 181 LLLIQLVLSCFSDRSPLFSETIHDPNPCPSSASFLSRITFWITGLIVRGVROPLEGSD 240  
QY 241 LWSLNKEDTSQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQPKSSKYDANSEVBAL 300  
Db 241 LWSLNKEDTSQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQPKSSKYDANSEVBAL 300



Db 241 LWSLNKEDTSEQVPELVKNNKCECAKTRKQPKVYSSKDPAPKESKSDVANEVEAL 300  
 Qy 301 IVKSPKEMNSLFKVLKTKTPPELMSFFPKAHLDMFSGPOLIKLLIKFVNDTKAPD 360  
 Db 301 IVKSPKEMNSLFKVLKTKTPPELMSFFPKAHLDMFSGPOLIKLLIKFVNDTKAPD 360  
 Qy 361 WQGYFYTVLLFTVACLOTLVHLQYFHCIVFSGMRKTAIVGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFTVACLOTLVHLQYFHCIVFSGMRKTAIVGAVYRKALVITNSARKSSTV 420  
 Qy 421 GEIVNLSVDAQRFMDLATYINIMWSAPLOVILLYLLNMLNLSVLAGVAVMLVPVN 480  
 Db 421 GEIVNLSVDAQRFMDLATYINIMWSAPLOVILLYLLNMLNLSVLAGVAVMLVPVN 480  
 Qy 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIRQBELKVLK 540  
 Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIRQBELKVLK 540  
 Qy 541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTTIDENNILDAQTAFAVSLALFNILRPPLNLP 600  
 Db 541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTTIDENNILDAQTAFAVSLALFNILRPPLNLP 600  
 Qy 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGTNSITVRNATFTWARSDDPT 660  
 Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGTNSITVRNATFTWARSDDPT 660  
 Qy 661 LMGITFSPGALVAVGVGCGKSLLSALLAEMDKVEGHVIAIKGSVAVYVPOQAWIQND 720  
 Db 661 LMGITFSPGALVAVGVGCGKSLLSALLAEMDKVEGHVIAIKGSVAVYVPOQAWIQND 720  
 Qy 721 SLRENILFGCOLLEPYYSVIOACALLPDLEILPSGDRTEIGEGKVNLSGGQKQVSLAR 780  
 Db 721 SLRENILFGCOLLEPYYSVIOACALLPDLEILPSGDRTEIGEGKVNLSGGQKQVSLAR 780  
 Qy 781 AVYSNADIYLFDDPLSADVAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 Db 781 AVYSNADIYLFDDPLSADVAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 Qy 841 MSGGKISEMGSYQELLARDGAPAEFLRYASTEQEODAEENGVTGVSFGPKGKAKOMENG 900  
 Db 841 MSGGKISEMGSYQELLARDGAPAEFLRYASTEQEODAEENGVTGVSFGPKGKAKOMENG 900  
 Qy 901 LVTDGAGKQQLQSSSSSYSGDISRHHNSTAELQAEAKKEETWKLMEADKAQTQGVKL 960  
 Db 901 LVTDGAGKQQLQSSSSSYSGDISRHHNSTAELQAEAKKEETWKLMEADKAQTQGVKL 960  
 Qy 961 SVYDYMKAIGLFIISFLPMCHVNSALASNYWLSLWTDPIVNGTQEHKVRLSVYG 1020  
 Db 961 SVYDYMKAIGLFIISFLPMCHVNSALASNYWLSLWTDPIVNGTQEHKVRLSVYG 1020  
 Qy 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVLDLHLSILRSPMSFFERTPSGNLVNRFKSEL 1080  
 Db 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVLDLHLSILRSPMSFFERTPSGNLVNRFKSEL 1080  
 Qy 1081 DTVDSMIPEVIMKFMGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 Db 1081 DTVDSMIPEVIMKFMGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 Qy 1141 KLESVSRSPTYSHFNETILLGVSVIRAFEEQERFTHQSDLKVDENQKAYPSIVANRWLA 1200  
 Db 1141 KLESVSRSPTYSHFNETILLGVSVIRAFEEQERFTHQSDLKVDENQKAYPSIVANRWLA 1200  
 Qy 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
 Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
 Qy 1261 VERLKEYSETEKAPWQIETAPPSPWPQGVREFRNYCLARYREDLDFVLRHINVTINGG 1320  
 Db 1261 VERLKEYSETEKAPWQIETAPPSPWPQGVREFRNYCLARYREDLDFVLRHINVTINGG 1320  
 Qy 1321 EXKVGIVGRTGAGKSLTGLFRINESARGEIIIDGINIAKIGLHDLRFXKIIIPQDPVLF 1380  
 Db 1321 EXKVGIVGRTGAGKSLTGLFRINESARGEIIIDGINIAKIGLHDLRFXKIIIPQDPVLF 1380

Qy 1381 SGSLRMNLDPFSQYSDSEVWTSLELAHLKDFVSALPKLDHECAEGGENLSVGORQLVCL 1440  
 Db 1381 SGSLRMNLDPFSQYSDSEVWTSLELAHLKDFVSALPKLDHECAEGGENLSVGORQLVCL 1440  
 Qy 1441 ARALLRKTILVLDEATAAVDLETTDDLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTILVLDEATAAVDLETTDDLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Qy 1501 DKGEIOEYGAPSDILOQORGLFYSNAXKDAGLV 1531  
 Db 1501 DKGEIOEYGAPSDILOQORGLFYSNAXKDAGLV 1531

Search completed: March 18, 2005, 11:02:13  
 Job time : 147.282 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:20 ; Search time 45.2651 Seconds  
(without alignments)  
4096.085 Million cell updates/sec

Title: US-10-665-283-1  
Perfect score: 9903  
Sequence: 1 MALRGFCSDGSDPLMDWNV.....RSVAVAKPKFSIPDLSL 1927

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7849	79.3	1531	1 DVHUAR	multidrug resistan
2	4484.5	45.3	1527	2 JE0336	canalicular multis
3	3587	36.2	1545	1 S71841	multidrug resistan
4	3507.5	35.4	1541	1 S71839	canalicular multid
5	3289	33.2	1494	2 E89447	protein F57C12.4 [
6	3225	32.6	1573	2 T21219	hypothetical prote
7	3195.5	32.3	1502	2 T42216	multidrug resistan
8	2938	28.7	1515	1 S51863	cadmium resistance
9	2732	27.6	1398	2 T20434	hypothetical prote
10	2548	25.7	1478	2 T39712	ABC transporter SP
11	2384.5	24.1	1623	2 T01369	ABC transporter At
12	2356.5	23.8	1622	2 D86428	glutathione S-conj
13	2326	23.5	1495	2 E86428	probable ABC trans
14	2312	23.3	1144	2 T27408	hypothetical prote
15	2294.5	23.2	1559	1 S64757	probable membrane
16	2263.5	22.9	1488	2 F86428	probable ABC trans
17	2238.5	22.6	1516	2 F84919	glutathione-conjug
18	2230	22.5	1539	2 T48059	ABC transporter-li
19	2144.5	21.7	1355	2 T00961	hypothetical prote
20	2144.5	21.7	1514	2 T52080	multi resistance p
21	2140	21.6	1515	2 T52081	MRP-like ABC trans
22	2131.5	21.5	1490	2 T47840	multi resistance p
23	2129	21.5	1545	2 T46645	sulfonylurea recep
24	2123	21.4	1153	2 T26883	hypothetical prote
25	2109	21.3	1545	2 T42751	sulfonylurea recep
26	2104.5	21.3	1511	2 T42711	sulfonylurea recep
27	2098	21.2	1546	2 T42728	sulfonylurea recep
28	2069.5	20.9	1389	2 T47796	ABC transporter-li
29	2062.5	20.8	1661	2 S64800	probable membrane

ALIGNMENTS

RESULT 1

DVHUAR

multidrug resistance protein (cell line H69AR) - human  
N/Alternate names: multidrug resistance-associated protein (MRP)  
C/Species: Homo sapiens (man)  
C/Date: 30-Jun-1993 #sequence\_revision 05-Dec-1998 #text\_change 19-Jan-2001  
C/Accession: A44231; A37495  
R/Cole, S.P.C.; Bhargwaj, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almqvist, K.C.;  
Science 258, 1650-1654, 1992  
A/Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer  
A/Reference number: A44231; MUID:93088080; PMID:1360704  
A/Accession: A44231  
A/Status: nucleic acid sequence not shown  
A/Molecule type: mRNA  
A/Residues: 'MAPTSGTMSRGIPATPSPAFRTSSCGCLVFTSGPV', 50-1531 <CO1>  
A/Cross-references: GB:L05628; NID:g1835658  
A/Experimental source: small cell lung carcinoma cell line H69AR  
A/Note: sequence extracted from NCBI backbone (NCBIP:119851); this sequence has been corrected  
R/Cole, S.P.C.; Deeley, R.G.  
Science 260, 879, 1993  
A/Title: Multidrug resistance-associated protein: sequence correction.  
A/Reference number: A37495; MUID:93262415; PMID:8098549  
A/Accession: A37495  
A/Status: not compared with conceptual translation  
A/Molecule type: mRNA  
A/Residues: 1-80 <CO2>  
A/Cross-references: GB:L05628; NID:g1835658  
A/Note: sequence extracted from NCBI backbone (NCBIP:131929)  
C/Genetics:  
A/Gene: GDB:MRP  
A/Cross-references: GDB:136335; OMIM:158343  
A/Map position: 16p13.1-16p13.1  
C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
C/Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transmembrane  
F:661-844/Domain: ATP-binding cassette homology <ABC1>  
F:678-685/Region: nucleotide-binding motif A (P-loop)  
F:788-792/Region: nucleotide-binding motif B  
F:1310-1503/Domain: ATP-binding cassette homology <ABC2>  
F:1327-1334/Region: nucleotide-binding motif A (P-loop)  
F:1450-1454/Region: nucleotide-binding motif B

Query Match 79.3% ; Score 7849; DB 1; Length 1531;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSDGSDPLMDWNVTNTSNPDFTKCFQNTVLVWVPCFYLMACFPFFYLYLSRH 60  
Db 1 MALRGFCSDGSDPLMDWNVTNTSNPDFTKCFQNTVLVWVPCFYLMACFPFFYLYLSRH 60  
QY 61 DRGIQMTPLNKTALGFLWVWVPCFYLMACFPFFYLYLSRH 120  
Db 61 DRGIQMTPLNKTALGFLWVWVPCFYLMACFPFFYLYLSRH 120

QY 121 TLLIQLERRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Db 121 TLLIQLERRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 QY 181 LLLIQLVLSGFCSDRSPFLSETIHDNPNCPSSASFLSRITFWIITGLIVRGYRQPLEGSD 240  
 Db 181 LLLIQLVLSGFCSDRSPFLSETIHDNPNCPSSASFLSRITFWIITGLIVRGYRQPLEGSD 240  
 QY 241 LMSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVVYSSKDPAPKPKSSKVDANEVEAL 300  
 Db 241 LMSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVVYSSKDPAPKPKSSKVDANEVEAL 300  
 QY 301 IVKSPKKNWPSLFKVLKXKTFPGPYFLMSFFFAIHDLMFSGPQILKLIKEVNDTKAPD 360  
 Db 301 IVKSPKKNWPSLFKVLKXKTFPGPYFLMSFFFAIHDLMFSGPQILKLIKEVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTAQLTLVLHQYFHCIFVSGMRKTAIVGAYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTAQLTLVLHQYFHCIFVSGMRKTAIVGAYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINNIWSAPLOVILALYLLWNLGSPVLGAVMVLMPVN 480  
 Db 421 GEIVNLSVDAQRFMDLATYINNIWSAPLOVILALYLLWNLGSPVLGAVMVLMPVN 480  
 QY 481 AVAMKTKTYQVAHMKSKDNRIKLMEILNGIKVLKYAWELAFKDKVLATROBELKVLK 540  
 Db 481 AVAMKTKTYQVAHMKSKDNRIKLMEILNGIKVLKYAWELAFKDKVLATROBELKVLK 540  
 QY 541 KSAYLSAVGTFTWVCTPFLVALCTFAVYVTDENNILDAQTAFAVSLALFNILRFPNLILP 600  
 Db 541 KSAYLSAVGTFTWVCTPFLVALCTFAVYVTDENNILDAQTAFAVSLALFNILRFPNLILP 600  
 QY 601 WYISSIVQASVSLKRLRFLSHEELEPDSIERRPVKDGGTNSIIVRNATFTWARSDDPT 660  
 Db 601 WYISSIVQASVSLKRLRFLSHEELEPDSIERRPVKDGGTNSIIVRNATFTWARSDDPT 660  
 QY 661 LNGITFTSPEGALVAVGVGGGCKSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIOND 720  
 Db 661 LNGITFTSPEGALVAVGVGGGCKSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIOND 720  
 QY 721 SURENILFGCOLLEPYRSVIOACALLPDLPLPSGDRTEIGEKNVLSGGQKQVSLAR 780  
 Db 721 SURENILFGCOLLEPYRSVIOACALLPDLPLPSGDRTEIGEKNVLSGGQKQVSLAR 780  
 QY 781 AVYSNADYLFDDPLSADVAHVGHKIFENVIGPKGMLKKNKTRILVTHSMSYLPQVDVILV 840  
 Db 781 AVYSNADYLFDDPLSADVAHVGHKIFENVIGPKGMLKKNKTRILVTHSMSYLPQVDVILV 840  
 QY 841 MSGGKISEMGYSQELLARDGAFAPFLRTYASTEQQDAEENGVTGVSFGPKGKAKOMENG 900  
 Db 841 MSGGKISEMGYSQELLARDGAFAPFLRTYASTEQQDAEENGVTGVSFGPKGKAKOMENG 900  
 QY 901 LVTDAGLQOLQOLSSSSYSYSGDISRHNSHTAELOKAEAKKEETWKLMEADKAQTQVKL 960  
 Db 901 LVTDAGLQOLQOLSSSSYSYSGDISRHNSHTAELOKAEAKKEETWKLMEADKAQTQVKL 960  
 QY 961 SVYDYMKAIGLFIISFLSIFELFMCNHNVSALASNYWLSLWTDPIVNGTQETHKVLRSYVG 1020  
 Db 961 SVYDYMKAIGLFIISFLSIFELFMCNHNVSALASNYWLSLWTDPIVNGTQETHKVLRSYVG 1020  
 QY 1021 ALGISQGIADVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSEL 1080  
 Db 1021 ALGISQGIADVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSEL 1080  
 QY 1081 DTVDMSIPEVIMFMGSLFNIVIGACIVILLATPIAAIIIPBLGIYFPVQFYVASSRQL 1140  
 Db 1081 DTVDMSIPEVIMFMGSLFNIVIGACIVILLATPIAAIIIPBLGIYFPVQFYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLVKDNOKAYYPSIVANRWLA 1200  
 Db 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLVKDNOKAYYPSIVANRWLA 1200

QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVYSLOVTTYLNLWVRMSEMETNIVA 1260  
 Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVYSLOVTTYLNLWVRMSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKEAPWQIQETAPPSSWPQGVFRPNRYCLAVREDLDFVLRHINVTINGG 1320  
 Db 1261 VERLKEYSETEKEAPWQIQETAPPSSWPQGVFRPNRYCLAVREDLDFVLRHINVTINGG 1320  
 QY 1321 EKVGIIVGTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITIIPQDPVLF 1380  
 Db 1321 EKVGIIVGTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITIIPQDPVLF 1380  
 QY 1381 SGLSRMNLDPSSQYSDREVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 Db 1381 SGLSRMNLDPSSQYSDREVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 QY 1441 ARALLRKTIIILVDEATAAVALDETDLLIQSTIRIQFEDCTVLTIAHRLNTIMDYTRVIL 1500  
 Db 1441 ARALLRKTIIILVDEATAAVALDETDLLIQSTIRIQFEDCTVLTIAHRLNTIMDYTRVIL 1500  
 QY 1501 DKGEIOEYGAPESSDILQORGLFYSMAGLVL 1531  
 Db 1501 DKGEIOEYGAPESSDILQORGLFYSMAGLVL 1531  
 RESULT 2  
 JEO336  
 canalicular multispecific organic anion transporter - human  
 C:Species: Homo sapiens (man)  
 C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 09-Jul-2004  
 C:Accession: JEO336  
 R:Uchiumi, T.; Hinoshita, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Toh, S.; Furukawa, M.;  
 Biochem. Biophys. Res. Commun. 252, 103-110, 1998  
 A:Title: Isolation of a novel human canalicular multispecific organic anion transporter,  
 t.  
 A:Reference number: JEO336; MUID:99032812; PMID:9813153  
 A:Accession: JEO336  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1527 <UCH>  
 A:Cross-references: UNIPROT:O15438; GB:AF083552  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
 P:Keywords: ATP  
 P:1306-1499/Domain: ATP-binding cassette homology <ABC2>  
 Query Match 45.3%; Score 4484.5; DB 2; Length 1527;  
 Best Local Similarity 56.6%; Pred. No. 1.1e-294;  
 Matches 871; Conservative 261; Mismatches 372; Indels 35; Gaps 9;  
 QY 8 SADGSDPLWDMNVTWNTSNPDTKCFQNTVLVWVPCFVLWACFPFYFLYLSRHRDRGIOM 67  
 Db 7 SGEIGSKFWDSNLSVHTENFDLTFCFQNSLLWVPCYILWVALPCYLLYLRRHRCRGVIL 66  
 QY 68 TPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPFVFLVSPTLGTTTLATLFIQLE 127  
 Db 67 SHLSKLXWLVGLVLCVSWADLFYSPHGLVHGRAPAFVFFVPLVGVVMTLLATLIQYE 126  
 QY 128 RRKQVQSGMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVFSLLILQLV 187  
 Db 127 RLQGVQSGVLIIFWFLCVCAIIVPFRSKILLAKAEISDPFRFTFYIHFALVLSALI 186  
 QY 188 LSCFSDKSPFLSETIHDNPNCPSSASFLSRITFWIITGLIVRGYRQPLEGSDLSNKE 247  
 Db 187 LACFREPFPFSAKXVDNPNPETSAGLSRLRFLFWMTKAIYGYRHPLEKDLWSLKEE 246  
 QY 248 DTSQVVPVVLVKNWKECAKTRKQPVVYSSKDPAPKPKSSKVDANEVEALIVKSPQK 307  
 Db 247 DRSQMVQQLLEAWRKQEKQTAH-----KASAPGK-----NAGGEVLLGARPRP 294  
 QY 308 EWNPSLVKVLKXKTFPGPYFLMSFFFAIHDLMFSGPQILKLIKEVNDTKAPDQGVYFT 367  
 Db 295 R-KPSFLKALLATFGSSFLISACFKLIQDLSLFINPOLLSILIRFISPNPMSGWMGLVA 353

368 VLLFTVACLOTLVLHOFYHFCVFGMRIRKTAIVAGVYRKALVITNSARKSSTVGBIVNLM 427  
 354 GLMFLCSMMQSLILOHYHYIFVTGKFRGTGIMGVYRKALVITNSVKEASTVGBIVNLM 413  
 428 SVDAQREMDLATYINNIWASAPLOVILALVLLNINLGPVSLAGVAVNVLMPVNAVWAKMT 487  
 414 SVDAQREMDLAPFLNLLWASAPLOIILAIYFLNQNIGPSVLAVFVWLLDPLINGAVAVOK 473  
 488 KTYQVAHMKSKDNRIKLMEIINGIKVLKLYAWELAFKDKVLAIROEELKVLKKSAYLSA 547  
 474 RAPQVKQMKLKDRIKLMEIINGIKVLKLYAWELAFKDKVLAIROEELKVLKKSAYLSA 533  
 548 VGTFTWCTPPLVALCTFAYVYITIBENNLDQAQFVSLALFNILRPLNLIPLMVISIV 607  
 534 TTTFTWMCSPFLVTLITLWVYVDPNNVLDKAKAFVSVSLFNILRPLNLIPLMVISIV 593  
 608 QASVSLKRLRIFLSHEELPDSEIRRPVKDGGTNSITVRNATFTWARSDDPTLNGITFS 667  
 594 QASVSLKRIQQFLSQBELDQSPVERKTSIFG---YAITHSCTFTWAQDLPTLHSLDIQ 650  
 668 IPEGALVAVVGVGCGSKSLLSALLAEMDKVEGHVAIKGSVAVVPPQAWIQNDSLENI 727  
 651 VPKGALVAVVGVGCGSKSLLSALLAEMDKVEGHVAIKGSVAVVPPQAWIQNCTIQENVL 710  
 728 FCQCLEEPYRSVIOACALLPDLEILPQSDRTEIGEKGWNLGGQKQKQVSLARAVYNAD 787  
 711 FGKALNPKEYQOTLEACALLADLEMLPGGDQTEIGEKGWNLGGQKQKQVSLARAVYNAD 770  
 788 IYLFDDPLSADVAHVKGKHFENVIPKGMKXKTRILVTHSMYSLPOVDVILVMGGKTS 847  
 771 IFLLDDPLSADVAHVKGKHFENVIPKGMKXKTRILVTHSMYSLPOVDVILVMGGKTS 830  
 848 EMGYSQELLARDGAFELRTYASTEQEDQABENGVTVGSGPCKEAKQWENGML----- 901  
 831 EMGYPALLORNGSFANFLCNAPDEDDQGHLEDSWTALEGAEDKEALLIEDTLNHTDLT 890  
 902 ----VTDSAGKQLOKOLSSSSSYSGD-----ISRHHNSTAE-LQXAKKEETWKLMEAD 951  
 891 DNDPVTYVYVQKQFMRLSALS--DGEQGQRPVRRRLHGPSEKVPQTEAKADGA--LTQBE 947  
 952 KAQTQGVKLSVYWDYKATGLFISFLSIFELMCNHVSALASNYMLSLMTDDPVLVNGTQEH 1011  
 948 KAAIGTVELSVFWDYKATGLFISFLSIFELMCNHVSALASNYMLSLMTDDPVLVNGTQEH 1007  
 1012 TKVRLSVYALGISOQIAVFGSMVAVSIGGILASRLHVDLHLSILRSPMSFFERTPSPG 1071  
 1008 TSLRLGVYALGISOQIAVFGSMVAVSIGGILASRLHVDLHLSILRSPMSFFERTPSPG 1067  
 1072 LVNRSKELDTVDSIPEVIMKPMGSLFNIVGACIVILLATPIAIIIPPLGLIYFFVOR 1131  
 1068 ILNCFSKDIYVDEVILAPVILMLNSFFNAISTLVVIMASTPLFTVILPLAVLYTLVOR 1127  
 1132 FYVASSRQLKLESVRSRVPVSHENETLGVSVIRAFEEQERFIHQSDLKVDENOKAYP 1191  
 1128 FYAATSRQKRLSVRSRVPVSHENETLGVSVIRAFEEQERFIHQSDLKVDENOKAYP 1187  
 1192 SIVANRWLAVRLEVCNGLVPAALFAVIRSRHSLAGLVGLSVSYSLQVTTYNLNLVRMS 1251  
 1188 YIISNRWLSIGVEFGNVCVLPALFAVIRSRHSLAGLVGLSVSYSLQVTTYNLNLVRMS 1247  
 1252 SEMETNIVAVRLEKEYSETEKAPQIOETAPSSWPQVGRVEFRNYCLRYREDLOFVL 1311  
 1248 SDLESNIVAVRLEKEYSETEKAPQIOETAPSSWPQVGRVEFRNYCLRYREDLOFVL 1307  
 1312 HINVTINGEKVIGVIRTCAGKSSLTGLFRINESAGEIITDGINAKIGHDLRFKIT 1371  
 1308 DLSLHVHGKGVIGVIRTCAGKSSLTGLFRINESAGEIITDGINAKIGHDLRFKIT 1367  
 1372 IIPQDPVPLSGSLRMLNLPFSQVDEEVWTSLELAHLKDFVSALPKDLKHECAEGENLS 1431  
 1368 IIPQDPVPLSGSLRMLNLPFSQVDEEVWTSLELAHLKDFVSALPKDLKHECAEGENLS 1427  
 1432 VQORQLVCLARALKRKTILVLDATAAVDLETDLLIOSTIRTPQEDCTVLTIAHRLNTI 1491

Db 1428 VQORQLVCLARALKRKTILVLDATAAVDLETDLLIOSTIRTPQEDCTVLTIAHRLNTI 1487  
 QY 1492 MDYTRVIVLDKGEIYGAGPSDLLQOORGLFYSMAXDAGL 1530  
 Db 1488 MDYTRVIVLDKGVVAEFDSPANLIAARGIFYGMARDAGL 1526  
 RESULT 3  
 S71841  
 multidrug resistance protein, canalicular - human  
 C:Species: Homo sapiens (man)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 19-Jan-2001  
 A:Accession: S71841, S71840  
 R:Koenig, J.; Keppeler, D.  
 submitted to the EMBL Data Library, August 1996  
 A:Reference number: S71841  
 A:Accession: S71841  
 A:Molecule type: mRNA  
 A:Residues: 1-1545 <KOE>  
 A:Cross-references: EMBL:X96395; NID:gi507819; PIDN:CAA65259.1; PID:gi514568  
 R:Buechler, M.; Koenig, J.; Brom, M.; Kartenbeck, J.; Spring, H.; Horie, T.; Keppeler, D.  
 J. Biol. Chem. 271, 15091-15098, 1996  
 A:Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance  
 A:Reference number: S71839; MUID:96279006; PMID:8662992  
 A:Accession: S71840  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1415-1429 'VP', 1432-1455 'B', 1457-1545 <BUE>  
 A:Cross-references: EMBL:X96395  
 C:Genetics:  
 A:Gene: GDB:ABCC2; CMOAT; ABC; MRP2; CNRP; DJS  
 A:Cross-references: GDB:6089489; OMIM:601107  
 A:Map position: 10q24-10q24  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette protein  
 F:654-837/Domain: ATP-binding; nucleotide binding; P-loop; transmembrane protein  
 F:671-678/Region: ATP-binding; nucleotide-binding motif A (P-loop)  
 F:1317-1510/Domain: ATP-binding cassette homology <ABC2>  
 F:1334-1341/Region: nucleotide-binding motif A (P-loop)  
 Query Match 36.2%; Score 3587; DB 1; Length 1545;  
 Best Local Similarity 47.5%; Pred. No. 6.8e-234;  
 Matches 746; Conservative 286; Mismatches 461; Indels 78; Gaps 20;  
 QY 3 LRGCASDGSDPLMDWNVNTS-----NPDTKCFQNTLVVWPCFYLWACFPFVFL-- 55  
 Db 2 LEKFCN-----STFWNSSFLDSEADPLCFEQTLVWILGFLMLLAPQLLHV 51  
 QY 56 YLSRHDRGYIQWTPNKTKTALGFLMWIVCWADLPYSFWERSRGIFLAPVFLVSPFLGI 115  
 Db 52 YKSRTKRSSTTKLYLAK-QVFGVFL--LLAAIETALVLTEDSGQATVPAVRYTNPSTLY-L 108  
 QY 116 TLLATFLIQLERRKGVQSSGIMLT-FWLVALVCLALILRSKIMTALKEDAQVDLFRDIT 174  
 Db 109 GTWLVLVLIQYRQWCVQKNSWFLSFLWILSILCGTFQFTLIRTLQGD-NSNLAYSCL 167  
 QY 175 FVYVPSLLLIQLVLSFCSDRSPFSETIHDNPNCPSESSASFLSRITFWMITGLIVRGYRQ 234  
 Db 168 FFISYGFQILLIIFSAFSEN-----NESSNPSSIASFLSSITYSYWDSIILKGYKR 219  
 QY 235 PLEGSDLWSLNKEDTSEQVVPVLVKNWKECAKTRK-----QPVKVYVSSKD 281  
 Db 220 PLTLEDVWEVDEBEMTKTLVSKFETHMKRELQKARRALQRRQEKSSQNSGARGLPGLNKN 279  
 QY 282 PAQPKESKVDANEEVEALIVKSPQKWNPS--LFKVLKYKTPGYPFLMSFFPKAHLHLM 339  
 Db 280 QSQSDALV---EDVEKKKKSGTKQVPSKWLKALFKTFYVWLLSKSLKLNDIFT 336  
 QY 340 FSGPQILKLLIKFVNDRTPADPWQGYFTVLLFVTAQCLQTLVHLQYFHCFCVSGMRKTA 399  
 Db 337 FVSPQLLKLLISFASDRDTYLMIGYLCALLFTALIIQSFCLQCFCFLGKGVKRTAI 396



```

Db      344  GFKVSSNSYVWFYICAILMFAVTLIQSFCLQSYFQHCFLVGCVRVTTVMSSYIKKALT 403
Qy      411  TNSARKSSTVGIIVNLSVDAQFMDLATYINNIWSAPLOVTLALYLWNLGSPVLAVG 470
Db      404  SNLARKQYITGETVNLMSVDSQKMDATNMQVLWSSVIQITLSIFFLRELGPSILAGV 463
Qy      471  AVNVLMPVNAVMAKTKTYQVAHMKSKDNRIKLAMEILNGIKVLKYLAWELAFKDKVLA 530
Db      464  GWVLLIPVNGVLATKIRNIQONMKNDKRLKIMNEILSGIKILYFAWEPFQEQVG 523
Qy      531  IQEELKVLKKSAYLSAVGCTFWCTPFLVALCTFAVYVITIDENNILDAQTAFLVALPN 590
Db      524  IRKELKNLRFQLOSLIFILQITPILVSVTFVSVYVLDVDSANVLAEKAFSTITLNF 583
Qy      591  ILRFPNLIPMVISSTVOASVSLKRLIFLSHEELEPDSIERPPVKDGGTNSITVRNAT 650
Db      584  ILRFPMLPMTVSTTLOASVSDRLERYLGGDDLDTSAIRVSNPD----KAVFSPAS 639
Qy      651  FTWARSDDPTLNGITPFIPEGALVAVGVQGGKSKSLASALLAEKMDKVEGHVAIKGSVAY 710
Db      640  FTWDPLEATIQVNLDIKPGQLVAVGVGVGSKSLVSAMLGEMENVHGHITIOGSTAY 699
Qy      711  VPQAWIQNDSURENIFCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKNVLSG 770
Db      700  VPQSWIQNGTIKDNILFSGSEYNEKYYQVLKACALLPDLLEILPGDMAEIGEKNLSG 759
Qy      771  GOKORVSLARAVYSNADIVLFDPLSANDAHVGHIFENVIGPKMLKNTKILYTHSMS 830
Db      760  GOKORVSLARAAQADIVYLDPLSANDAHVGHIFENVIGPKMLKNTKILYTHGHI 819
Qy      831  YLPQVDVITVMSGGKISEMGYSQELLARDGAFELRTYAS-TEOQDAEENGVTGVSGP 889
Db      820  FLPOVDEIVLKGITLKGYSYRDLDDKGVFARNWKTFFMKHSGPGEATVNN----- 872
Qy      890  GKEAKQMGEMGLVT-----DSAG-----KOLQOLSSSSSYSGDISRHENSTABLQ-- 935
Db      873  DSEAEDDDGLIPTMBEIPEDAASLAMRRNSLRRTLSRSSSSRRRGKSLNKLKNV 932
Qy      936  ---KAEAKKEETKLEADKATQGVKLSVYDYNKAIGLFTSLFIFCMCHVYSALAS 992
Db      933  NVLKEKEVEQOKLTKKEFEVGTGVKFSIYLYKIQAVGWMSILFIILPGLUNVAFGS 992
Qy      993  NYWLSLWT--DDPIVNGT---QHETKVLRSVYGALGTSQIAVFGYSMAVSGGILASRCL 1048
Db      993  NLWLSAWTSDNLNGTNNSSSHRDWRIGVFGALGLAQICLLISTLSIYACRNASKAL 1052
Qy      1049  HVDLLHSILRSPMSFFPTPSGNLVNRFSEKELDTVDMSIPEVIMKPMGSLFNVIACIVI 1108
Db      1053  HQQLLTNILRAPMRFDDTPTGRIVNRFSGDISTVDDLLPQTLRSWMMCFGIAGTLVMI 1112
Qy      1109  LLATPIAAIIIPGLIYFFQRFYVYASSROLKRLSVRSRSPVSHFNETLGVSVIRAF 1168
Db      1113  CMATPVFAIIIPLSLITSVQVYVATSRQRLRSDSVTKSPIYSHFSTVGLPIIRAF 1172
Qy      1169  EQERFIHOSDLKVDENKAYPSIVANRWLVRLECVGNCIVLFAALFAVISRHSLSAG 1228
Db      1173  EQQRFPLANWEKQIDINOKCVSWITSNFWLRLVGNLVVFCALLVIYRKLTCGD 1232
Qy      1229  LVGLSVSYSLQVTTYLNWLVMSSEMETNIVAVERLKEYSETKEXAPWQIOETAPSSWP 1288
Db      1233  VVGFLSNALNITQTLNWLVRMTSEATNIVAVERISEYINVENEAPM-VTKRPPADWP 1291
Qy      1289  QVGRVEFRNYCLRYREDLDFVLRHINVTNGEKGIVGRTGAGKSSLTGLFRINESAE 1348
Db      1292  RHGEIOFNQYRYPPELDELKGTCTCNIKSGEKVGWGTGAGKSSLNCLFRIUESAG 1351
Qy      1349  GEIIDGINIAKGLHDLRFKTIIPQDPVLFSGLSRMMLDPPFSQVSDDEVWTSLELAHL 1408
Db      1352  GOIIDGIDVASGLHDLRERLTIIPQDPILFSGSLRMNLDLPNKYSDEVWRALELAHL 1411
Qy      1409  KDFVSLPDKLDHECAGGENISVGORQVLVCLARALLRKTILVLDEATAAVDLETDDLI 1468

```

```

Db      1412  RSFVSGLOGLLSEVTEGGDNLSTGQROLLCLGRVLRKSKILVLDENTAAVDLETDSLI 1471
Qy      1469  OSTRTQPEDCTVLTIAHRLANTIMDYTRVILVDKGEIOEYGAPSDLLQORGLFYMAKDA 1528
Db      1472  QTTIRKESQCTVITIAHRLTIMDSKIMWLDNGKIYVYGSPELLSNRGSFYLMAKEA 1531
Qy      1529  GL 1530
Db      1532  GI 1533

RESULT 5
E89447
protein F57C12.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: E89447
R:anonymouse, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: E89447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1494 <STO>
A:Cross-references: UNIPROT:Q20943; GB:chr_X; PIDN:AAA83299.1; PID:g1118071; GSPDB:GN060
C:Genetics:
A:Map position: X
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match      33.2%; Score 3289; DB 2; Length 1494;
Best Local Similarity 47.6%; Pred. No. 9,7e-214;
Matches 696; Conservative 239; Mismatches 435; Indels 92; Gaps 18;

Qy      119  LATLIQIERRKGVSSGIMLTFWLVLCALAILRSLKIMTALKEDAQVDLFRDITFYVY 178
Db      76  LALILTACKNKGITTSVITLYLLVWVGIPFPRFYLGGFIYNEALEGIRATLYILA 135
Qy      179  FSLLIQLVLCSPDRSPLFSETHDHPNCPSESSAFSLRITFWITGLIVRGYRQPLEG 238
Db      136  FTFSALEFLCCFAD---VPSDMYKSESCPEYATSFNRLTFQWFTGLAYLGNKKSLEN 192
Qy      239  SDSLNLKEDSEQVVPVLVKNWKECAKTRQPKVYVSSKDP--AOPKESKVDANEE 296
Db      193  EDLMDLNEIDKAEINLIPSMQNLKPRIDEYH-QNIK-----KPSAALPKN----- 237
Qy      297  VEALIVKSPQKEWNPSPFKVLYKTFGPYFLMSFFFKAIHDLMMFSGPQILKLLIKFVNDT 356
Db      238  -----HPSFVPIPTKYTYLLAGFFYKLCFDMQLQFLAPOLLKQLIGFIEDK 284
Qy      357  KAPDQGVFYTVLLPVTACLOTLVHQYHCTFCVSGMKIKTAVIGAVRKALVITNSARK 416
Db      285  NOPWVGISGVIMFSSFLQSMFLHQYHSMFRLGMHVRSLTSAVYSKALNSNEARK 344
Qy      417  SSTGEIVNLMSVDAQRFMDLATYINMLTWSAPLOVILALYLWNLGSPVLAVGVMLM 476
Db      345  GKTIGAIVNLMSVDLQIKQDMAPTMLFWSAPLQIFLSIYFLWKFLGVAALAGLVVLLA 404
Qy      477  VPVNAVMAKTKTYQVAHMKSKDNRIKLMEINLNGIKVLKYLAWELAFKDKVLAIRQEEL 536
Db      405  LPVNGLIAIQMKCKQTEQMKLKDRIKMSBILNGMKVLKYSWERSMENMVKIREREL 464
Qy      537  KVLKKSAYLSAVGTTTWTCTPFLVALCTFAVYVITD-ENNILDAQTAFAVSLAFILRFP 595
Db      465  HILKLSYFMAAIVFESWICAPFLASVIFVYVYLDPENNVLTPEITFVALSFDLFRMP 524
Qy      596  LNILPMWISSIVQAQSVSLKRLRIFLSHEELEPDSIERRPVKGGGTNSITVRNATFWAR 655
Db      525  LANVAVYGEAVQCVSVNTRLKEFFAAEMSPQT-----SISGETDTSALEVNGFLSWSS 580

```







887 -RPRSDAAPVKGS-----TSEAQMPEPSLDDVEVTGLTAGEDSVQYGRVKSATYLSYLA 940  
 970 IGLFISFLSFLFCMCHVGHASLNTYLSLWTDPIVNGTOETHKVRSLVYGLGTSQGT 1029  
 941 VGTPLCTYTLFLFLCQVASFQGYWLSUAWDDPVVDGQKMSALRGSIFGLGCLQIG 1000  
 1030 VFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRFSKELDTVDSMIPE 1089  
 1001 LFSMAAVPLGGRASCLLFRSLLDVAVSPGFFERTPVGNLLNRFSKETDIDVDIPD 1060  
 1090 VIKPMGSLFNIVIGACIVILLATPIAAIIIPGLIYFVQRYVASSRQLRLESVRS 1149  
 1061 KMTLLTYAFGLLEGLAVSMATPLAIVAILPLMLLYAGFQSLYVATCCQLRLESASYS 1120  
 1150 PVSFHNETHLLGVSVIRAPPEOERFTHQSDLVADENOKAYPSIVANRWLAVRLECVNC 1209  
 1121 SVCSHLAETFGSQVVRPAQGPPTAQHDALMDENQRISFRLVADRWLANLELLGNG 1180  
 1210 IYVLAALFAVIRHRSLSAGLVLSYSLSQVTTYLNVLVRMSSEMETNIVAVERLKEYSE 1269  
 1181 LVFVAATCAVLSKAHLSAGLAFSVSAALQVTTQLQWVRSWTDLENSVAVERQDVYH 1240  
 1270 TEKEAPWQIOETAPSSPQVQVRPNVCLRYREDLDFVLRHINVTNGEKGIVGRT 1329  
 1241 TPKEAPWRLPSSAAQPLWPCGQIEFRDFGLRHPPELPMVAVQVSLKIHAKEKVGIVGT 1300  
 1330 GAGKSSLTGLPRINESAGEITIIDGINIAKIGLHDLRFKITIIPQDPVLPSGSLRMNLD 1389  
 1301 GAGKSSLTGWLRLQEGTEGGIWDGVFTDMLGLTLRSRTIIPQDPVLPSGSLRMNLD 1360  
 1390 PFSQYSDEEVTSLBLAKDFVSALPKDLHECAEGGENLSVGQRLVCLARALLRRTK 1449  
 1361 LIQENTDEGIWALETVOLKAFVTSPLGQLQVCEGSGQDLSVGQKQLCLARALLRRTQ 1420  
 1450 IIVLDEATAAVDLEDDLIQSTIRQFREDCTVLTIAHLNIMDYTRVIVLDKGBIQBYG 1509  
 1421 IIVLDEATASVDPGTEIQWQALERWFAQCTVLLIAHLRLSVYMCARVLVMDQVABSG 1480  
 1510 APSDLLQORGLFYSMAXDAGL 1530  
 1481 SPAQLAQKQFYRLAQESGL 1501

RESULT 8  
 S51863  
 cadmium resistance protein YCF1 - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein YD9302.11c; protein YDR135c  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 09-Jul-2004  
 R:Accession: S51863; A55352; S50233  
 R:Oliver, K.; Harris, D.  
 submitted to the EMBL Data Library, February 1995  
 A:Reference number: S51853  
 A:Accession: S51863  
 A:Molecule type: DNA  
 A:Residues: 1-1515 <OLI>  
 A:Cross-references: UNIPROT:P39109; EMBL:Z48179; NID:G665657; PIDN:CAA88217.1; PID:G6656  
 R:Szyzypka, M.S.; Wemmie, J.A.; Moya-Rowley, W.S.; Thiele, D.J.  
 J. Biol. Chem. 269, 22853-22857, 1994  
 A:Title: A yeast metal resistance protein similar to human cystic fibrosis transmembrane  
 A:Reference number: A55352; MUID:94357936; PMID:7521334  
 A:Accession: A55352  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-679, 'R', 681-1515 <SZC>  
 A:Cross-references: GB:L35237; NID:G556464; PIDN:AAA50353.1; PID:G556465  
 C:Genetics:  
 A:Gene: SGD:YCF1; MIPS:YDR135c  
 A:Cross-references: SGD:S0002542; MIPS:YDR135c  
 A:Map position: 4R  
 C:Function:  
 A:Description: required for cadmium resistance  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein; yeast vacuole  
 F:287-308/Domain: transmembrane #status predicted <TM1>  
 F:345-366/Domain: transmembrane #status predicted <TM2>  
 F:421-442/Domain: transmembrane #status predicted <TM3>  
 F:446-467/Domain: transmembrane #status predicted <TM4>  
 F:534-555/Domain: transmembrane #status predicted <TM5>  
 F:558-580/Domain: transmembrane #status predicted <TM6>  
 F:646-829/Domain: ATP-binding cassette homology <ABC1>  
 F:663-670/Region: nucleotide-binding motif A (P-loop)  
 F:951-972/Domain: transmembrane #status predicted <TM7>  
 F:995-1016/Domain: transmembrane #status predicted <TM8>  
 F:1068-1088/Domain: transmembrane #status predicted <TM9>  
 F:1092-1113/Domain: transmembrane #status predicted <TM10>  
 F:1179-1200/Domain: transmembrane #status predicted <TM11>  
 F:1208-1229/Domain: transmembrane #status predicted <TM12>  
 F:1289-1483/Domain: ATP-binding cassette homology <ABC2>  
 F:1306-1313/Region: nucleotide-binding motif A (P-loop)

Query Match 28.7%; Score 2838; DB 1; Length 1515;

Best Local Similarity 40.3%; Pred. No. 3.4e-183;  
 Matches 632; Conservative 295; Mismatches 495; Indels 146; Gaps 27;

QY 28 DFTKCFQNTVLVWPCFYLWACFPFYF-----LYLSRHD-----RGYIQMTPLNKT 75  
 DB 27 DFTQCFIDGVIL-----NLSAIFMTTIGIRDVNLCKKHSGIKYRRNWIIVSRMALVLL 81  
 QY 76 ALGFLMTVWCADILFYSEWRSRGIFLAPVPLVSPITLIGITTLATFLIQLERRKGVSS 135  
 DB 82 EIAF-----VSLASNLISKEEAEEN---FTIVSQYASTMLSLFVALAHWIEYDR--SVVAN 132  
 QY 136 GIMLTFTVLVALCALATILRSKIMTALKEDAQVDLFRDITFVYVFS-----LLLQLVLSC 190  
 DB 133 TVLIFYMLFETFGNFAKILNIRHYEG-----IWSQGTGFTLITFQVITC 180  
 QY 191 PS-----DRSPL-----FSETHDNPCEPSSASFLSRITFTWITGLIVRGYRQPLEG 238  
 DB 181 ASILLLEALPKLPHQHIIHQTLTRKPNPYDSANIFSRITFTSWMGLMKTGYEKYLVE 240  
 QY 239 SDLSLWLNKEDTSEOVVPLVKNWKECAKTRKQPVVYSSKDPAPKSESSKVDANEVE 298  
 DB 241 ADLYLPRNFSSEELSOKLEKNWENEL-----KQKS----- 271  
 QY 299 ALIVKPSQKEMNPSLFLVYKTFGPFYLMGFFFKAIHDLMMFSGPQLIKLKEFVND--- 355  
 DB 272 -----NPSLSWAICRTFGSKMLAAFAFKAIHDVLAFTQPLRLILKEVTDVNS 320  
 QY 356 -----TKAPDWOGYFYTVLLFTVACQLTILVHQYPHICFVSGMRKITA 398  
 DB 321 ERQDHSLSQGFENNHPQKLPIVRGFLIAFAMFLVGTQTSVLHQYFLNVENTGMYKSA 380  
 QY 399 VIGAVYRKALVITNSARKSSTVGBIVNLSMDAQRFDMLATYINMWSAPLQVILALYLL 458  
 DB 381 LTALIYOKSLVLSNEASGLSTGDIVNLSMDVQKLDQLTQWLNLWSGPPQIILCLYSL 440  
 QY 459 WNLGSPSLVAGVAVVLMVPMVNAVMAKTKTYQVAHMSKDNRIKLNELINGIKVLKLY 518  
 DB 441 YKLIGNSMWVGVIILVIMPLNSPLMRIOKKLQKSKMYKXDRTRVISEILNLSKLY 500  
 QY 519 AWELAFKDKVLAIR-QBELKVLKSAVLSAVGTFTVCTPFLVALCTFAVYVTTDENNIL 577  
 DB 501 AWEKPYREKLEEVNKNELKNTKLGCYMAVTSQFQNVLPVLSVSCCTFAVVFV-YTEDRAL 559  
 QY 578 DAQTAFFVSLAFNLIRFPLNIPMWISSIVOASVSLKRLRIFLSHEELEPDSIRRRPVKD 637  
 DB 560 TTDLVFFALTLNLLSPFLMIIPVNLNSPIEASVIGRLFTFTFNEELQDSDVORLPKVK 619  
 QY 638 GGGTNSITV-RNATFTTWARSD--PPTLNGITFTSIPEGALVAVVGQVCGKSSLSALLAE 694  
 DB 620 NIGDVAINIGDATFLWQRPVKVVALKNINFOAKKGNLTCIVGVKSGKTKALLSCMLGD 679  
 QY 695 MDKVEGHVAKGSVAVYVPOQAWIQNDLSRLNIFGCOLBEPYRYSVTOACALLPDLILP 754  
 DB 680 LFRVKGPAVHGSVAVYVSPVWIMNGTVKENILFGHRYDAEFYEKTIKACALITDILM 739



755 SGRTEIGKGVNLSGGQKQVSLARAVYSNADIVLFDPLSAVDHVGKHIFENVIGPK 814  
 740 DGDGKTAVGKGLSGGQKARLSARAVARADTYLDDPLAADVHVARHLEHVLGN 799  
 815 GMLKNTKRLVTHSHSYLPOVDVIVMSGGKISEMGSOELLA-RDGAFAEFLRYASTE 873  
 800 GLHTTKTKVLATNKVSALSADSIALLDNGEITQOQTYDEITKADSPMLKLLNNYK-- 857  
 874 QBDAAENGVTGSGPGKAKQWENGMLVTDGAGKQORQLSSSSSYSGDISHHNSTAE 933  
 858 -----XNKGKSEFPGDSSSESVRESIPV-EGELQOLKLNLDLPGNSDAISLRASDAT 911  
 934 LOKAEAKKEETWKLMEADKAQTKGVKSLVYDYMKA-----ICLFISFLSIFLFCNHYVS 988  
 912 LGSIDFGDDE--NIAKREHREGQKVKWNIYLVAKACPKSCVCFILFIVISMF-- 964  
 989 ALASNYWLSLWTDVFNCTQHTKVLRSVYGALGISQGIAGVGYSMASVI--GGILASRC 1047  
 965 SYMGNNVWLKHWSEVNSRYSGSNPNAAYLAIFYALGIGSALATLIQTIIVLWVFCTIHASKY 1024  
 1048 LHVLDLLHSILRSPMGFFERTPGNLVNRFSKELDTVDSMIPEVIMFGMSLFPNVIGACTV 1107  
 1025 LNLWNTSVLRAPMTFFETPIGRILNRFNSDIYKVDALLGRTFSQFFVNAVKVTFITIV 1084  
 1108 ILLATPIAAIIIPPLGLIYFFQRFVYVASSRQLKRLSVSPVYSHFNETLLGVSVIRA 1167  
 1085 ICATTWQFIIFIIPLSVFYIYQVYVYRTSRRLRLDSITRSPYISHFOETLGLLATVRG 1144  
 1168 FEOERFTHQSDLKVDENOKAYPSIVANRWLAIVLECVGNCIVLFAALFAV--LSRHS 1225  
 1145 YSQQRFSHINOQRIDNNNSAFYPSINANRWLAIVLEIGSIILGAATLSVFLKQGT 1204  
 1226 SAGLVLSYSYSLQVTVYVNLVNRSEMETNIVAVERLKEYSETEKAPMOIQETAPPS 1285  
 1205 TAGWGLSLYALQITQTLNVRMTVEVETNIVSVERIKEFADLKSEAPLIVEGHRPPK 1264  
 1286 SWPOGRVFRNYCLRYREDLDFVLRHNVITNGGKGVIGRTGAGKSSLTGLFRINE 1345  
 1265 EWPSQGDIFKNYSTRYRPELOLVKLHINHIKPNKGVIGRTGAGKSSLTALFRMIE 1324  
 1346 SAGEHIIIDGINIAKGLHDLRPFKTIIPQDPVLSGSLRMLNDPFSQVSEVWTSLEL 1405  
 1325 ASEGNIVIDNIAINEIGLYDLRHKLSTIIPQDSQVFEGTVRENDIPNQYTDIAIWRAL 1384  
 1406 AHLKDFVSLP--DKLDHECAEGENLSVGORQLVCLARALLRKTILVLDEATAVDLET 1464  
 1385 SHLKEHVLMSNDGLDAQTEGGGNSLVGQRLCLLARMLVPSKILVLDEATAVDVET 1444  
 1465 DDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPEDDL--QORGLFYS 1523  
 1445 DRVQETIRTAFKDRTILTIAHRLNTIMDSRIIVLDNGKVAEPFSPGQLLSDNKSIFY 1504  
 1524 MAKDAGLV 1531  
 1505 LCMWEAGLV 1512

RESULT 9  
 Hypothetical protein E03G2.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T20434  
 R:McMurray, A.  
 submitted to the EMBL Data Library, November 1995  
 A:Reference number: Z19275  
 A:Accession: T20434  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1398 <WIL>  
 A:Cross-references: UNIPROT:Q19048; EMBL:Z68113; PIDN:CAA92148.1; GSPDB:GN000028; CESP:E0  
 A:Experimental source: clone E03G2

C:Genetics:  
 A:Gene: CESP:E03G2.2  
 A:Map position: X  
 A:Introns: 73/3; 113/3; 176/1; 216/1; 258/2; 384/3; 470/3; 549/1; 644/3; 757/3; 764/2; 8  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 27.6%; Score 2732; DB 2; Length 1398;  
 Best Local Similarity 41.1%; Pred. No. 4.5e-176;  
 Matches 591; Conservative 287; Mismatches 472; Indels 88; Gaps 19;

QY 115 ITTLATLTIQLERKGVSSGIMLTFWLVALVCALAILRSKIMTALKED--AQVDLRED 172  
 DB 27 LALLTIATNEV---RRAGIHSSGPLFCIWMFLFAVAAPPEFYQMMTGTSGSELPVARDIFPY 83  
 QY 173 ITFYVYFLLLIQVLSCFSDRSPLFSETIHDPN-----PCPESSASFLSRITFTFWMITGL 227  
 DB 84 VAYLTYFFLVVAEFLVHFVSDPFPM-----PRGYNLKCPEENANFISRQLLLWFTQI 136  
 QY 228 IVRGYRQPLEGSDMLSLNKEDTSEOVVPLVKNWKKCEAKTRKQPKV-------VYSS 279  
 DB 137 ISLGYVERTLVADVDPEMDSQMDQEYLKARWKTWMLKQTEKAREKQVLDKRRERARTGSE 196  
 QY 280 KDPAPQKESKVDAN--EEVEALIVKSPQKWNPSLFKLYKTPGYFLMSFFFAIHDL 338  
 DB 197 KAPLLGTNNYGANVLDDKDRVIVQ-----PSVITLQIMKWEILGGGFIKFLSLL 249  
 QY 339 MFSGQIILKLIKFVNDTKAPDMQGYFYTVLLFVTAQLQTLVHLQYFHCYFVSGMRKITA 398  
 DB 250 QFANPTFLNYLLFIETPNAPLINGIGLAVGLFAGQIKSLFMNTYFIAMTRVGAKIQTM 309  
 QY 399 VIGAYRKALVITNSARKSTVGEIIVNLSMDVDAQRFMDLATYVNNIMWSAPLOVIALYLL 458  
 DB 310 LSCAVYEKSLLSNTARRERTVGMENVILSDVDRFRMITPQIQOYQWSSPFQIICMVL 369  
 QY 459 WNLGSPVLGAVVAVMLVNPVNAVAMKTKYOVAAHMSKNRIKLMNEILNGIKVLKLY 518  
 DB 370 SOTIGVAVAGIVVMISIVPINICVSMITKRWQLKMKYKDBERILNEVLNGIKVVKLS 429  
 QY 519 AWEIAFKDKVLAIROEELKVLKKSAYLSAVGTFTWVCTPFLVALCTFAVYVITDENNIID 578  
 DB 430 AWEIAMEETIERVRDKELKMKIKQSALLKTFADCLNVGAPVFVALSSFTVFVLLDPKVL 489  
 QY 579 AQTAFVSLALFNILRFPNLPIMVITSSIVQASVSLKRLRIFLISHBELSPDSIERPVDKG 638  
 DB 490 PNIAFVSLFNLRLGPLMAAEELVAQTVQLVSNKRVRTFLCEKEVDTAADKE--IRGE 548  
 QY 639 GGTNSITVRNATFTWASDDPPTLNGITPISIEGALVAVVGVQVCCCKSSLSALLAEMDKV 598  
 DB 549 LYTNVTEIHSGSFAWDSAEARILSDIEFLAGSKELTVTVVSGSKSLLLAALGEMEKV 608  
 QY 699 EGHVAIKGSVAVPQQAWIQNDSLRNIFLFGQLEEPYVRSVIQACALLPDLIELPSGDR 758  
 DB 609 CGYGVGRSVAYLSQPPWILNQLKKNVLMQADLNDVLYKKVIESCALKEDLKQLPDGDD 668  
 QY 759 TEIGEGVNLGGQKQVRSARVSNADIVLFDPLSAVDHVGKHIFENVIGPKMLK 818  
 DB 669 TEIGEGINLSGGQKARIALARAVYQSKDVYFLDDPLSAVDHVGKHIFENVIGPNGMLS 728  
 QY 819 NKTRILVTHMSYLPQVDVIVMSGCKTISEMGSYQELLARQGAFAEFLTYASTQEODA 878  
 DB 729 HTTRILVTNCTSFQESGKIIVMKG-----ETRYESSG 761  
 QY 879 EENGVTGSGPGKAKQAKQWENGMLVTDGAGKQORQLSSSSSYSGDISHHNSTAE 938  
 DB 762 EESG-----GEENSDILPGSIASGSRMSLSR--LSK-----ISRKSKSSIVEK-- 804  
 QY 939 AKKETWKLMEADKAQTKGVKLSVYDYMKAIGLPIFLSIFLFCNHYVS--ALASNYWLS 997  
 DB 805 -KKPDA--LITKEEAAIGRVNPGVYLLYFKAMGIVTVYVLAIAVAVLVNVSFALGRSLMT 861  
 QY 998 LWTDDPVPNGTQEHKTV--RLSVYCALGISQGIAGVGYSMASVIGGILASRCILHVDLHS 1055  
 DB 862 AWSNDANIDINHPDNTMSVGARLGVAGFGITVIFLFFSLVLLIGGVAASKNLHKLPLHN 921



QY 1227 AGLVGLSVSYSLQVTTYLNWLVMSSEMETNIVAVRLKEYSETKEAPWQIQETAPSS 1286  
Db 1173 SGLVGLSLSYAVQIQSTLTVVRQSDVETNIVSVERMLVIGLSEAPSIIPDRPPEG 1232  
QY 1287 WQVGRVFRNYCLYREDLDVLRHINTVINGGKVGIVGTGAGKSLTGLPRINES 1346  
Db 1233 WPSGAIRFDHYSVRYENLPLVANDISVNIKPBQKIGIVGTGAGKSLTGLPRLLIEP 1292  
QY 1347 ABGEIIDIIGINIAGIKGLHDLREPKITIPDPVLFSGSLRMNLDPPFSQYSDBEVWTSLELA 1406  
Db 1293 TSGDQLDDINITSGLHDLRSRLAIIPQENCAFEGTIRENLDPNANATDEIWHALEAA 1352  
QY 1407 HLKDFVSALPKDLDEHCAEGENLVSQGRQLVCLARALLRKTILVDEATAAVDLETD 1466  
Db 1353 SLKQFIQTLDDGLYSLRVTEGGANLSSGQRLMCLTRALLTPRVLLDEATAAVDETDA 1412  
QY 1467 LIQSTIRQFEDCTVLTIAHRLNTIMDYTRVILDKGEIOEYCAPSDILQOR-GLFYDMA 1525  
Db 1413 IVQRTIRERFNDRTILTIAHRLNTIMDYTRVILDKGEIOEYCAPSDILQOR-GLFYDMA 1472  
QY 1526 KDAGLV 1531  
Db 1473 KESGLI 1478

RESULT 11  
T01369  
ABC transporter AtMRP2 [imported] - Arabidopsis thaliana  
N:Alternate names: multidrug-resistance protein homolog T29F13.13  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: T01369; D84759  
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
submitted to the EMBL Data Library, May 1998  
A:Description: Arabidopsis thaliana chromosome II BAC T29F13 genomic sequence.  
A:Reference number: Z14179  
A:Accession: T01369  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1623 <ROU>  
A:Cross-references: UNIPROT:064590; EMBL:AC003096; NID:g3132469; PID:g3132479  
A:Experimental source: cultivar Columbia  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.;  
eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: D84759  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1623 <STO>  
A:Cross-references: GB:AE002093; NID:g3132479; PIDN:AA016268.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g34660; T29F13.13  
A:Map position: 2  
A:Introns: 110/3; 170/2; 183/3; 236/1; 274/2; 301/2; 309/3; 328/3; 350/3; 375/3; 411/3;  
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
F:1259-1452/Domain: ATP-binding cassette homology <ABC>

Query Match 24.1%; Score 2384.5; DB 2; Length 1623;  
Best Local Similarity 35.2%; Pred. No. 1.9e-152;  
Matches 561; Conservative 274; Mismatches 534; Indels 223; Gaps 30;

QY 44 FYLVACFPF-YFVLSRHRDGYIQMTPLNKTATG-FLVWCVWADLFYSFW----- 94  
Db 5 FIEWYCKPVNGVWTKQVAFAYPCATDSVLGISQLVLLVLC-----LYRIMWALKD 60  
QY 95 -----ERSRGIFLA-----PVFLVSPITLLGTTL-----LATLIQLERK 130  
Db 61 HKVERFLSRNLNYFLALLAAVATPELFRL---IMGISVLDGFGGLPPP----- 109  
QY 131 GVQSSGIMLTFWLVAVCALAIRSKIMTALKEDAQVLDLFRDITFYVYFS-----LL 182

Db 110 --EAFGLGVKAFANGAVMMILMETKI-----YIRELRWYRFVAVIYALVGMV 156  
QY 183 LIQLVLS-----CFSDRSPLSETIHD--- 204  
Db 157 LLNLVLSVKEYSSVVLVLYTSEVCAQVLFGLLPMHLPLNDTPGYNMPVSETVDDVEY 216  
QY 205 -----NPNCPESASFLSRITFWITGLIVRGYRQPLEGSDLSLNKEDSEQVVPVLV 258  
Db 217 EEISDGOQICPEKHANIPDKIFFSWMNPLMTLGSKRPLTEKDVWVYLDTDOTETLFTSFQ 276  
QY 259 KWKKECAKTRKQPKVYVSSKDPAPKESKVDANBEVEALIVKSPKWNPSLPKVLV 318  
Db 277 HSWDKELQK-----PO-----PWLRLALN 295  
QY 319 KTFPGPYFLMSFFKAIHDLMMFSGPOILKLIKLVNDTKAPDQWQGYFYTVLLFVTAQLQT 378  
Db 296 NSLGRFWMGFWKIGNDSCQFVGPLLLNQLLKSQOE-DAPAWGYIYAFSIFVGVVGV 354  
QY 379 LVHLQYFHCIFVSGNRKIRKTAIVAGYRKALVITNSARKSSTVGEIVNLSMDVAQRFMDLA 438  
Db 355 LCEAQYFQNVVRVGYRLRSALIAAVFRKSLRLTNEGRKFKQTGKITNLMTTDAESLQOIC 414  
QY 439 TYINMINSAPLOVILALYLLMLNLGSPVLGAVVAVMLAVPUNAVNAMKTKTYQVAHMSK 498  
Db 415 QSLHTMSAPFRIIIALILLYQQLGVASLIGALLLVLMFPPLQTVIISKQKLTKEGLQT 474  
QY 499 DNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROEELKVLKKSAYLSAVGTFTWCVTPF 558  
Db 475 DKRIGLMEVLAAMDVTKCYAWENSFQSKVQTVRDELDFSWFKSOLLGALNFIINSIPV 534  
QY 559 LVALCTFAVYVITDENNIDLAQTAQFVSLALFNILRPLNIPMLVPMVSISSIVQASVSKRL-- 616  
Db 535 LVTIVSGVFTLLGGD--LTPARAFTSLSLFAVLRFPFLMPLNIIITQVNVANVSLKRLSE 592  
QY 617 -----RIFLSHEELEPDSIERRPVKDGGGTNSIIVRNATFTW-ARSDPTPLNGITFSIP 669  
Db 593 VLATBERILLPNPIEP-----GPAISIRNGYFSWDSKGRDPTLSNINLNDVP 640  
QY 670 EGALVAVVGVGVCGKSSLSALLAEMDKV-BGHVAIKGSVAVVPOQAQWIONDSLRENILF 728  
Db 641 LGLSLVAVVSGTGEKTSLSAILGELPATSDAIVTLRGSVAVVPOVSWIFNATVRDNLIF 700  
QY 729 GCQLEPYRYSVIOACALLPDLLEILPSGDRTEIGSGVNLGSGQKQKORVSLARAVYSNADI 788  
Db 701 GSPFDREKVERAIDVTSKHDLELLPGGDLTEIGERGYNISGGQKQKORVSMARAVYSNSDV 760  
QY 789 YLFDDPLSAVDAHVGKHFENVIGPKGMLKNKTRILVTHSMYSYLPOVDVILVMSGKILSE 848  
Db 761 YIFDDPLSALDAHVGQVFEKCI--KRELQKTRVLVTNQLHFLSQVDRIVLVHEGTVKE 818  
QY 849 MGSYQELLARDGAFABEFLRTVASTEOQDAEENGVTGVSQPKGE--AKQMENGMLVTDSEA 906  
Db 819 EGTYEELSSNGLPFLORLMENAGKVEEY--SENGEAEADQTAEPQVANGNTNGLQMDGSD 876  
QY 907 KQQLORQLSSSSSYSGDISRHHNSTAEIQKAEKKEETWKLMEADKAQQTQGVKLISYWDY 966  
Db 877 DKK-----SKENKKGKSVLIKQESERETGVVSWRVLKRY 911  
QY 967 MKAI-GLFISFLSIFLFCNHNVSALSNYLSLWTDPIVNGT-QEHTKVRLS-VYGALG 1023  
Db 912 QDALGGAVVWMLLCYVLTVEFVRYTSSTLSSEWTD---AGTPKSHGLFLFNLYIALLS 967  
QY 1024 ISQGIAGVGSYMAVSGIGGILASRCLHVDLLHLSILRSPMSFPBTPSGNLVNRFSKELDTV 1083  
Db 968 FGQVLTUTNSTWLTNMSLSIAAKKLDHNDMLHLSILRAPMSFFHTNPLGRINRFAKDLGDI 1027  
QY 1084 DSMIEVKIMKMGSLFNIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQKRL 1143  
Db 1028 DRTVAVFVNMFMFGVQSLLSTVVLIGIVSTLSLWAIMPLLVLFYGAAYLYQNTAREVKRM 1087  
QY 1144 ESVSRSPVYSHNETLLGVSVIRAFEEQERFTHQSDLKVDENQKAYPIPSIVANRWLAVL 1203  
Db 1144 ESVSRSPVYSHNETLLGVSVIRAFEEQERFTHQSDLKVDENQKAYPIPSIVANRWLAVL 1203



Db 1232 PSQSGIKFEDVVRPELPVHLGVSPFLISPMQKVGIVGRTGAGKSSLLNALFRIVEL 1291  
 Qy 1348 EGEIIDGNIAGIKGLHDLRFKTTIIPQDPVLPFSGSLRMNLPFSGSYDEEVTSLSLAH 1407  
 Db 1292 KGRILLDECDIGFGLMDLRKVLGIIPQAPVLPFSGTVRNLDPFSEHNDADLWESLERAH 1351  
 Qy 1408 LKDFVSALPKDLHCEAGENLSVGQRLVCLARALKTKILVLDEATAAVIDLETDL 1467  
 Db 1352 LKDTIRNPLGLDAEVEAGENFSGQRLSLARALLRSKILVLDEATAAVIDRTDVL 1411  
 Qy 1468 IQSTIRTOEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYCAPSDLLQORGLFYSMAD 1527  
 Db 1412 IQKTIRPEFKSCMTLIIAHLNTIIDCDKVLVLDGKGVQFSPENLLNSGSSFSK--- 1468  
 Qy 1528 AGLVGGGGGMLSRKIIPEEYVLTRLAEDPAEPRYRTRE 1567  
 Db 1469 -----MVQSTGTANAELRSITLENK-----RTRE 1493  
  
 RESULT 13  
 E86428  
 probable ABC transporter [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: E86428  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 anen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: E86428  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1495 <STO>  
 A:Cross-references: UNIPROT:Q9C8H0; GB:AE005172; NID:g11055918; PIDN:AAG28288.1; GSPDB:C  
 C:Genetics:  
 A:Map position: 1  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
  
 Query Match 23.5%; Score 2326; DB 2; Length 1495;  
 Best Local Similarity 34.3%; Pred. No. 1.5e-148;  
 Matches 541; Conservative 314; Mismatches 566; Indels 154; Gaps 28;  
  
 Qy 2 ALRGFCS--ADGSDPLMDNWNVTWNTSNPDFTKCFQNTVLVWVP-----CFY-LWACPP 51  
 Db 5 ALNWYCKPVADGF-----WEKAVDGAAGATPCADISLVMLVSHFVLGLGCEYRIWIIP- 58  
 Qy 52 FYFLYLSRHRDGYQMTPLNKTALGFLILWVCWAD-----LFVSWERSRGIFLAPV 105  
 Db 59 -----HNTKAIQYVLRKKYKNCVLG-LIACVYVEPVLRVLMGSLDFMDDEETDFPPF 110  
 Qy 106 FLVSPILLGITVLTATFLITQLERKQVQSSGIMLTFWLVALVCLAILRSKIMTALKEDA 165  
 Db 111 EVASLMVEAFAMFSLMVLVIGLETQYKQVGFYVRFVGLVADAVLLDLVPLKNSIN 170  
 Qy 166 QVDLFRDIT---PYVYFSLILLIQLVLSCFSDRSPLFSEITHDP-----NPCCP 209  
 Db 171 RTALYLFISRCQALFGILLIYIPEL--DPYFGYHIVNPELDNVEYDALRGGEHICP 228  
 Qy 210 ESSASFLSRITFWITGLIVRGYRQPLEGSDLSLNKEDTSBOQVPLVKNMKCEAKTR 269  
 Db 229 ERHASIFSRITFWITGLIVRGYRQPLEGSDLSLNKEDTSBOQVPLVKNMKCEAKTR 269  
 Qy 270 KQPVKYVSSKDKPAQKESKVDANEVEALIVKSPQKWNPSLFKVLKTYTGPFLMSF 329  
 Db 287 -----PK-----PWLRLNLSLGGRFWLAG 307

Qy 330 FFKAIHDLMFSGPOILKLLIKFVNDTKAPDQGWYFYTVLLFVTAQLOTLVLHQYFHICF 389  
 Db 308 IFKIGNDLSQFVGPVILSHLLRSMQEGD-PAWGVGVYAFIIFGVVTLGLVCLCAQYFQNVW 366  
 Qy 390 VSGRIKTAIVGAVYRKALVITNSARKSSTVGEIIVNLSVDAQRPMDLATYINMWSAPL 449  
 Db 367 RVGFRLRSLTVAALPHKSLRLTHEARKNPAKQVNTMTTDDANALQOISQQUHLWSAPF 426  
 Qy 450 QVITLALYLLMLNLPGLSVLAGVAVMVLVPMVNAVMAMTKTYQVAHMKSKDNRIKLMNEIL 509  
 Db 427 RIIVSMILLYQOGLGVASLFGSILFLPLQTLIISKMKRLTKEGLQWTDKRVGITNEIL 486  
 Qy 510 NSIKVLKLYAYELAPKQVLAIRQBELKVLKKSAYLSAVGTFTWCTPPLVALCTPAPVYV 569  
 Db 487 SMDTVKCYAWKSPESRIGQIRNEELSFWFRKAQLLSAFNSFILNSIPVVVTVWSEGVFV 546  
 Qy 570 TIDENNIIDAQTAFAVSLALENLIRPLNLPMLVSIIVQASVSLKRL-RIFLSHEELEPD 628  
 Db 547 LIGGD--LTPARAFTSLSLFAVLRPLNMLNLLSQVNVANVSLQRIEELLSEERILAQ 604  
 Qy 629 SIERRPVKDGGSNTSITVRNATFTW-ARSDPPTLNGITFSIPEGALVAVVGQVGGCKSSL 687  
 Db 605 NPPLQP-----GTPAISIKNGYFSDSKTKPTLSDINLEIPVGLVAVVGTTGEGKTSL 659  
 Qy 688 LSALLAEMDKVE-GHVAIKGSVAYVYQQAQWIONDSLRNLPGLQLEBPYRSVIOACAL 746  
 Db 660 ISAMLGELSHAETTSWIRGSVAYVYQVSWIFNATVRENILFGSDPESERYWRAIDATAL 719  
 Qy 747 LPDLILPSGDETEIGEKNVLSGOKORVSLARAVYSNADYLPDDPLSAYDAHVGHKI 806  
 Db 720 QHDLPLPGRDLTEIGERGVNISGOKORVSNARAVYNSDVYIFDDPLSALDAHVAHQV 779  
 Qy 807 PENVIGPGMKLNKTRILVTHSMSYLPQVDVIVMSGGKISGMSGYQELLARDGAFAPFL 866  
 Db 780 FDSM--KDELGRKTRVLVTLQHLPLMDKILLVSEGMIKEGTFVELSKSGLFKKLM 837  
 Qy 867 RYASTEQEQDABENGVTGVSQPGKAKOMENGLVTDGAKQOLQRLSSSSSYSGDISR 926  
 Db 838 ENAGKMDATQEVNTN-----DENILKLGPTVTVDVS---ERNLGSTK----- 876  
 Qy 927 HNSTAEILQAKAEKEETWKLMEADKAQTVGKLVSYVDYMKAI-CLFISFLSIFLFCMN 985  
 Db 877 -----QGGKRRS-VLIQKBERETGIISNVLMRYKEAVGLVWMLLACYLAT 924  
 Qy 966 HVSALASNVLSLWTDTPVNGTQSHTK-----VRLSVYGALGISOGIAVFGYSMAVSIG 1040  
 Db 925 EVLRVSSSTWLSIWD-----QSTSKNYSFGFVIVVYVALLGFGQVAVFTNSFWLITS 977  
 Qy 1041 GILASRCLHVDLLHLSILRSPMSFFERTPSGNLVNRFPSKELDTVDSMPIVIMKFMGSLFN 1100  
 Db 978 SLHAARLHDAMLSSILRAPMLFFHTNPTGRVIRNRFSDIGIDIRNVANLMMNFMNLQW 1037  
 Qy 1101 VIGACIVILLATPIAAIIPPLGLIYFFVQRFVASSRQLKELSVSRSPVYSHNETLL 1160  
 Db 1038 LSTTALLTGVSTISLWAIMPLLIIFYAAYLYQSTSRVRELRDVSPTSPIAQFGEALN 1097  
 Qy 1161 GVSIVRAPEEQERFTHQSDLKVDENOKAYPSIVANRMLAVRLECVGNCIVLFAALFAVI 1220  
 Db 1098 GLSSIRAYKAYDRMAKINGSNDNIRPTLANTSSNRMLTIRLETGGVMIWLTATFAVL 1157  
 Qy 1221 SRHSLS-----AGLVGLSVSYSLQVTVYLNMLVRMSSEMETNIVAVERUKYSETEKEAP 1275  
 Db 1158 QNGTNNQAGFASTWGLLLSYTLNITSLLSGVLRQASRAENSLSNVSVERVGNIDLPSEAT 1217  
 Qy 1276 WOIQETAPPSSWPQVRVEFRNYCLRYREDLDFVLRHINVTINGEKKIVGRTGAGKSS 1335  
 Db 1218 DIENNRVPCWPGSGSIFKFDVHLRYRPLGLPGLHGLTFFVPSBEKVGVORTGAGKSS 1277  
 Qy 1336 LTLGLFRINESAGEIIDDGINIAKIGLHDLRFKTTIIPQDPVLPFSGSLRMNLDPPSQYS 1395  
 Db 1278 MLNALFRIVEVEKGRIMIDDCDVAKGLTDVRVLSIIPQSPVLFGSTVRFNIDPFSEHN 1337



Residues: 1-255 <BOX>".

Note: the authors translated the codon CAG for residue 248 as His  
Purnelle, B.; Goffeau, A.

Submitted to the EMBL Data Library, April 1996

Description: The sequence of 32 kb on the left arm of yeast chromosome XII reveals 14  
family and a new ABC transporter homologous to the human multidrug resistance protein.

Reference number: S69380

Accession: S69391

Molecule type: DNA

Residues: 1-1559 <P>

Cross-references: EMBL:X97560; NID:gl297003; PIDN:CAA66162.1; PID:gl297015

Mioega, T.; Zimmermann, F.K.

ast 12, 693-708, 1996

Title: Sequence analysis of the CEN12 region of *Saccharomyces cerevisiae* on a 43.7 kb  
conductance regulator protein CFTR.

Reference number: S70557; MUID:96405918; PMID:8810043

Accession: S70560

Status: nucleic acid sequence not shown; translation not shown

Molecule type: DNA

Residues: 1-1559 <MTW>

Residues: 1-1559 <MTW>

Cross-references: EMBL:X91488; NID:gl495203; PIDN:CAA62776.1; PID:gl495208

Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995

Genetics:

Gene: SGD-BPT1; MIPS: YLL015W

Cross-references: SGD:S0003938

Map position: 12L

Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
Keywords: ATP; nucleotide binding; P-loop; transmembrane protein

28-44/Domain: transmembrane #status predicted <TM1>

143-159/Domain: transmembrane #status predicted <TM2>

178-194/Domain: transmembrane #status predicted <TM3>

334-350/Domain: transmembrane #status predicted <TM4>

421-437/Domain: transmembrane #status predicted <TM5>

526-542/Domain: transmembrane #status predicted <TM6>

550-566/Domain: transmembrane #status predicted <TM7>

654-847/Domain: ATP-binding cassette homology <ABC1>

672-679/Region: nucleotide-binding motif A (P-loop)

974-990/Domain: transmembrane #status predicted <TM8>

1017-1033/Domain: transmembrane #status predicted <TM9>

1059-1115/Domain: transmembrane #status predicted <TM10>

1118-1134/Domain: transmembrane #status predicted <TM11>

1212-1228/Domain: transmembrane #status predicted <TM12>

1319-1529/Domain: ATP-binding cassette homology <ABC2>

1336-1343/Region: nucleotide-binding motif A (P-loop)

```
Db 1363 DGIDISDIGLFDLRSHLAIIPQDAQAFEGTVKTNLDPFNRYSEDEBKRAVEQAHKPHLE 1422
Qy 1414 AL-----PDKLDHECAEGGENLSVGQSQLVCLARALLRKTILVLD EA 1456
Db 1423 KMLHSPRGDDSNEDGNVNDILDVKINENGNSLVGQSQLCLARALINRSKILVLD EA 1482
Qy 1457 TAAVDLETDDLIOSTIRTOPECTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGA PSDLLQ 1516
Db 1483 TASVDMETDKIIQDTIRREFKORTILTIAHRTDVLDSDKIILVDOQGSVREFPSPKLS 1542
Qy 1517 QR-GLFYSMAXDAG 1529
Db 1543 DKTSIFYSLCEKGG 1556
```

Search completed: March 18, 2005, 11:05:27  
Job time : 62.2651 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:20 ; Search time 203.942 Seconds  
(without alignments)  
4838.524 Million cell updates/sec

Title: US-10-665-283-1  
Perfect score: 9903  
Sequence: 1 MALRGFCGADGSDPLWMDV.....RSVAVAKPKFSISPDLS 1927

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7860	79.4	1531	1 MRP1_HUMAN	P33527 homo sapien
2	7769	78.5	1515	2 Q9UQ99	Q9UQ99 homo sapien
3	7695	77.7	1531	2 Q864R9	Q864R9 macaca fasc
4	7689	77.6	1531	2 Q864S0	Q864S0 macaca fasc
5	7444	75.2	1459	2 Q9UQ97	Q9UQ97 homo sapien
6	7419.5	74.9	1456	2 Q6UR05	Q6UR05 canis fami
7	7272	73.4	1531	2 Q9UQA0	Q9UQA0 homo sapien
8	7230.5	73.0	1530	2 Q8HXQ5	Q8HXQ5 bos taurus
9	7104.5	71.7	1400	2 Q9UQ98	Q9UQ98 homo sapien
10	7002.5	70.7	1528	1 MRP1_MOUSE	O35379 mus musculu
11	6932.5	70.0	1532	2 Q810E4	Q810E4 rattus norv
12	6932.5	70.0	1532	2 Q8CG09	Q8CG09 rattus norv
13	6892	69.6	1523	2 Q810G9	Q810G9 rattus norv
14	6075	61.3	1215	2 Q68CP7	Q68CP7 homo sapien
15	4522	45.7	1519	2 Q80ZK8	Q80ZK8 mus musculu
16	4491.5	45.4	1527	1 MRP3_HUMAN	O15438 homo sapien
17	4422	44.7	1514	2 Q9G0A9	Q9G0A9 homo sapien
18	4396	44.4	1522	1 MRP3_RAT	Q88563 rattus norv
19	3952.5	39.9	1592	2 Q7Q1D7	Q7Q1D7 anopheles g
20	3885.5	39.2	1548	2 Q7KTC3	Q7KTC3 drosophila
21	3880.5	39.2	1548	2 Q7KTC1	Q7KTC1 drosophila
22	3879.5	39.2	1548	2 Q7KTB7	Q7KTB7 drosophila
23	3874.5	39.1	1548	2 Q7KTC2	Q7KTC2 drosophila
24	3860.5	39.0	1548	2 Q717N0	Q717N0 drosophila
25	3857	38.9	1549	2 Q7KTC0	Q7KTC0 drosophila
26	3845.5	38.8	1548	2 Q879C5	Q879C5 drosophila
27	3838.5	38.8	1548	2 Q7KTB8	Q7KTB8 drosophila
28	3836.5	38.7	1548	2 Q7KTB9	Q7KTB9 drosophila
29	3818.5	38.6	1548	2 Q7KTD0	Q7KTD0 drosophila
30	3813.5	38.5	1548	2 Q7KTC8	Q7KTC8 drosophila
31	3812.5	38.5	1548	2 Q7KTC4	Q7KTC4 drosophila

32	3807.5	38.4	1548	2 Q7KTC9	Q7KTC9 drosophila
33	3793.5	38.3	1548	2 Q7KTC7	Q7KTC7 drosophila
34	3790	38.3	1549	2 Q9VK56	Q9VK56 drosophila
35	3771.5	38.1	1548	2 Q7KTC5	Q7KTC5 drosophila
36	3769.5	38.1	1548	2 Q7KTC6	Q7KTC6 drosophila
37	3741	37.8	1564	2 Q8QG98	Q8QG98 raja erinac
38	3738.5	37.8	1567	2 Q6PH26	Q6PH26 brachydanio
39	3655	36.9	822	2 Q9JHS0	Q9JHS0 rattus norv
40	3653.5	36.9	1544	2 Q6PSM3	Q6PSM3 canis fami
41	3641.5	36.8	1544	2 Q9SM36	Q9SM36 canis fami
42	3625.5	36.6	796	2 Q8C7V6	Q8C7V6 mus musculu
43	3595	36.3	1545	1 MRP2_HUMAN	Q92887 homo sapien
44	3594	36.3	1544	2 Q9SL75	Q9SL75 macaca mula
45	3577	36.1	1564	1 MRP2_RABIT	Q28689 o canalicul

## ALIGNMENTS

RESULT 1  
MRP1\_HUMAN  
ID\_MRP1\_HUMAN STANDARD; PRT; 1531 AA.  
AC P33527; O14819; P78419;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DE Multidrug resistance-associated protein 1 (ATP-binding cassette, sub-family C, member 1).  
DE family C, member 1).  
GN Name=ABCC1; Synonyms=MRP, MRP1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9308080; PubMed=1360704;  
RA Cole S.P.C., Bhargava J.G., Gerlach J.H., Mackie J.E., Grant C.E., Almqvist K.C., Stewart A.J., Kurz E.U., Duncan A.M.V., Deeley R.G.;  
RT "Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line."  
RT Science 258:1650-1654(1992).  
RN [2]  
RP REVISIONS.  
RX MEDLINE=93262415; PubMed=8098549;  
RA Cole S.P.C., Deeley R.G.;  
RT "Multidrug resistance-associated protein: sequence correction."  
RL Science 260:879-879(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96251691; PubMed=8649356;  
RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C., Deeley R.G.;  
RT "Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter."  
RL Mol. Pharmacol. 49:962-971(1996).  
RN [4]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;  
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;  
RT "Analysis of the intron-exon organization of the human multidrug-resistance protein gene (MRP) and alternative splicing of its mRNA."  
RL Genomics 45:368-378(1997).  
RN [5]  
RP SEQUENCE OF 1131-1531 FROM N.A.  
RX MEDLINE=9942570; PubMed=10493829; DOI=10.1006/geno.1999.5927;  
RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R., Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;  
RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q."  
RL Genomics 60:295-308(1999).  
RN [6]

RP TOPOLOGY, AND CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=97442425; PubMed=9295302; DOI=10.1074/jbc.272.38.23623;  
RA Hipfner D.R., Almquist K.C., Leslie E.M., Gerlach J.H., Grant C.E.,  
RA Deeley R.G., Cole S.P.C.;  
RT "Membrane topology of the multidrug resistance protein (MRP). A study  
RT of glycosylation-site mutants reveals an extracytosolic NH2  
RT terminus";  
RL J. Biol. Chem. 272:23623-23630(1997).  
RN [7]

RP TOPOLOGY.  
RX MEDLINE=97476249; PubMed=93342225; DOI=10.1074/jbc.272.42.26479;  
RA Kast C., Gros P.;  
RT "Topology mapping of the amino-terminal half of multidrug resistance-  
RT associated protein by epitope insertion and immunofluorescence";  
RL J. Biol. Chem. 272:26479-26487(1997).  
RN [8]

RP TOPOLOGY.  
RX MEDLINE=98153110; PubMed=9485377; DOI=10.1021/bi972332v;  
RA Kast C., Gros P.;  
RT "Epitope insertion favors a six transmembrane domain model for the  
RT carboxy-terminal portion of the multidrug resistance-associated  
RT protein";  
RL Biochemistry 37:2305-2313(1998).  
RN [9]

RP MUTAGENESIS OF ASP-792; ASP-793; LYS-1333 AND 1454-ASP-GLU-1455.  
RX MEDLINE=21362977; PubMed=11469806; DOI=10.1006/abbi.2001.2441;  
RA Cui L., Hou Y.-X., Riordan J.R., Chang X.-B.;  
RT "Mutations of the Walker B motif in the first nucleotide binding  
RT domain of multidrug resistance protein MRP1 prevent conformational  
RT maturation";  
RL Arch. Biochem. Biophys. 392:153-161(2001).  
RN [10]

RP MUTAGENESIS OF TRP-1246  
RX MEDLINE=2123819; PubMed=11278867; DOI=10.1074/jbc.M011246200;  
RA Ito K., Olsen S.L., Qiu W., Deeley R.G., Cole S.P.C.;  
RT "Mutation of a single conserved tryptophan in multidrug resistance  
RT protein 1 (MRP1/ABCC1) results in loss of drug resistance and  
RT selective loss of organic anion transport";  
RL J. Biol. Chem. 276:15616-15624(2001).  
RN [11]

RP VARIANTS GLN-633 AND VAL-671.  
RX MEDLINE=20296630; PubMed=10835642; DOI=10.1038/76102;  
RA Le Saux O., Urban Z., Techuch C., Csizsar K., Bacchelli B.,  
RA Quagliari D., Pasquali-Ronchetti I., Pope F.M., Richards A., Terry S.,  
RA Bercovitch L., de Paape A., Boyd C.D.;  
RT "Mutations in a gene encoding an ABC transporter cause pseudoxanthoma  
RT elasticum";  
RL Nat. Genet. 25:223-227(2000).  
RN [12]

RP VARIANT VAL-671.  
RX MEDLINE=20283940; PubMed=10811882; DOI=10.1073/pnas.100041297;  
RA Ringfeil F., Lebowitz M.G., Christiano A.M., Utto J.;  
RT "Pseudoxanthoma elasticum: mutations in the MRP6 gene encoding a  
RT transmembrane ATP-binding cassette (ABC) transporter";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:6001-6006(2000).  
RN [13]

RP VARIANT SER-433, AND CHARACTERIZATION OF VARIANT VAL-671.  
RX MEDLINE=21578494; PubMed=11721885;  
RA Conrad S., Kauffmann H.-M., Ito K., Deeley R.G., Cole S.P.C.,  
RA Schrenk D.;  
RT "Identification of human multidrug resistance protein 1 (MRP1)  
RT mutations and characterization of a G671V substitution";  
RL J. Hum. Genet. 46:656-663(2001).  
RN [14]

RP VARIANTS MET-117 AND LEU-1512.  
RX MEDLINE=20579883; PubMed=11139250;  
RX DOI=10.1002/1098-1004(2001)17:1<74::AID-HUMU14>3.0.CO;2-F;  
RA Perdu J., Germain D.P.;  
RT "Identification of novel polymorphisms in the pms and MRP1 (ABCC1)  
RT genes at locus 16p13.1 and exclusion of both genes as responsible for  
RT pseudoxanthoma elasticum";  
RL Hum. Mutat. 17:74-75(2001).  
RN [15]

RP VARIANTS SER-43; ILE-73; GLN-723 AND GLN-1058.  
RX MEDLINE=21163848; PubMed=11266082;  
RA Ito S., Ieiri I., Tanabe M., Suzuki A., Higuchi S., Otsubo K.;  
RT "Polymorphism of the ABC transporter genes, MDR1, MRP1 and MRP2/CMOAT,  
RT in healthy Japanese subjects";  
RL Pharmacogenetics 11:175-184(2001).  
CC -!- FUNCTION: May participate directly in the active transport of  
CC drugs into subcellular organelles or influence drug distribution  
CC indirectly.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=8;  
CC Comment=Additional isoforms seem to exist. Experimental  
CC confirmation may be lacking for some isoforms;  
CC Name=All isoforms;  
CC IsoId=P33527-1; Sequence=Displayed;  
CC Name=Delexon-17;  
CC IsoId=P33527-2; Sequence=VSP\_000037;  
CC Name=Delexon-18;  
CC IsoId=P33527-3; Sequence=VSP\_000038;  
CC Name=Delexon-30;  
CC IsoId=P33527-4; Sequence=VSP\_000039;  
CC Name=Delexon-17-18;  
CC IsoId=P33527-5; Sequence=VSP\_000037, VSP\_000038;  
CC Name=Delexon-17-30;  
CC IsoId=P33527-6; Sequence=VSP\_000037, VSP\_000039;  
CC Name=Delexon-18-30;  
CC IsoId=P33527-7; Sequence=VSP\_000038, VSP\_000039;  
CC Name=Delexon-17-18-30;  
CC IsoId=P33527-8; Sequence=VSP\_000037, VSP\_000038, VSP\_000039;  
CC -!- TISSUE SPECIFICITY: Lung, testis and peripheral blood mononuclear  
CC cells.  
CC -!- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.  
CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;  
CC WWW="http://www.infobiogen.fr/services/chroncancer/Genes/MRPID106.html".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L05628; AAB46616.1; --  
CC EMBL; AF022853; AAB83983.1; JOINED.  
CC EMBL; AF022824; AAB83983.1; JOINED.  
CC EMBL; AF022825; AAB83983.1; JOINED.  
CC EMBL; AF022826; AAB83983.1; JOINED.  
CC EMBL; AF022827; AAB83983.1; JOINED.  
CC EMBL; AF022828; AAB83983.1; JOINED.  
CC EMBL; AF022829; AAB83983.1; JOINED.  
CC EMBL; AF022830; AAB83983.1; JOINED.  
CC EMBL; AF022831; AAB83983.1; JOINED.  
CC EMBL; AF022832; AAB83983.1; JOINED.  
CC EMBL; AF022833; AAB83983.1; JOINED.  
CC EMBL; AF022834; AAB83983.1; JOINED.  
CC EMBL; AF022835; AAB83983.1; JOINED.  
CC EMBL; AF022836; AAB83983.1; JOINED.  
CC EMBL; AF022837; AAB83983.1; JOINED.  
CC EMBL; AF022838; AAB83983.1; JOINED.  
CC EMBL; AF022839; AAB83983.1; JOINED.  
CC EMBL; AF022840; AAB83983.1; JOINED.  
CC EMBL; AF022841; AAB83983.1; JOINED.  
CC EMBL; AF022842; AAB83983.1; JOINED.  
CC EMBL; AF022843; AAB83983.1; JOINED.  
CC EMBL; AF022844; AAB83983.1; JOINED.  
CC EMBL; AF022845; AAB83983.1; JOINED.  
CC EMBL; AF022846; AAB83983.1; JOINED.  
CC EMBL; AF022847; AAB83983.1; JOINED.  
CC EMBL; AF022848; AAB83983.1; JOINED.  
CC EMBL; AF022849; AAB83983.1; JOINED.  
CC EMBL; AF022850; AAB83983.1; JOINED.

Db	961	SVYWDYKAIGLFI	SFLSIFLPCN	HVSALASNYW	LSLWTD	DDTVNGTQ	EH	TKV	LSVYVG	1021			
QY	1021	ALGISQGI	AVFGYSMAV	SIGGILAS	RCLHVD	LLHSILR	SPMSFF	TPSGN	LNVNRSKEL	1080			
Db	1021	ALGISQGI	AVFGYSMAV	SIGGILAS	RCLHVD	LLHSILR	SPMSFF	TPSGN	LNVNRSKEL	1080			
QY	1081	DTVDSMIP	EVIKWFMG	SLFNVTG	ACACIV	LLATPIA	IIIPPLGL	IIVFFQ	RVFVASSRQL	1140			
Db	1081	DTVDSMIP	EVIKWFMG	SLFNVTG	ACACIV	LLATPIA	IIIPPLGL	IIVFFQ	RVFVASSRQL	1140			
QY	1141	KRLSVRSR	SPVYSHF	NETLLG	SVIRAF	EEQERF	FIHQSD	LKVD	ENOKAYYP	SIVANRWLA	1200		
Db	1141	KRLSVRSR	SPVYSHF	NETLLG	SVIRAF	EEQERF	FIHQSD	LKVD	ENOKAYYP	SIVANRWLA	1200		
QY	1201	VRLECVGNC	IVLFAAL	FAVISRHS	LSAGLV	GVLSV	SYSLQ	VTTYL	NMLVRMSSE	MTNIVA	1260		
Db	1201	VRLECVGNC	IVLFAAL	FAVISRHS	LSAGLV	GVLSV	SYSLQ	VTTYL	NMLVRMSSE	MTNIVA	1260		
QY	1261	VERLKEYS	ETEKAP	WQIQET	APPSSP	QWQGR	VEFENY	CLRY	VEDLDF	VLRHINV	TVINGG	1320	
Db	1261	VERLKEYS	ETEKAP	WQIQET	APPSSP	QWQGR	VEFENY	CLRY	VEDLDF	VLRHINV	TVINGG	1320	
QY	1321	EKVGIVGR	TGAGK	SSLTGL	FRINESA	EGEIIID	GINIAK	IIGH	LHDLRF	KITIIIP	QDPVLF	1380	
Db	1321	EKVGIVGR	TGAGK	SSLTGL	FRINESA	EGEIIID	GINIAK	IIGH	LHDLRF	KITIIIP	QDPVLF	1380	
QY	1381	SGSLRN	NLDPFS	QYSDEE	VMTSLE	LAHLK	DFVSAL	PKLD	HECAEG	ENLSVG	QRLVCL	1440	
Db	1381	SGSLRN	NLDPFS	QYSDEE	VMTSLE	LAHLK	DFVSAL	PKLD	HECAEG	ENLSVG	QRLVCL	1440	
QY	1441	ARALLRKT	KLVLJ	DEATAV	DLTDD	LLIOST	ITHTQ	FEDCT	VLTI	AHRLNT	IMDYTR	VIVL	1500
Db	1441	ARALLRKT	KLVLJ	DEATAV	DLTDD	LLIOST	ITHTQ	FEDCT	VLTI	AHRLNT	IMDYTR	VIVL	1500
QY	1501	DKGEIO	EGYCAP	SDDLQ	QRG	LFYSMA	KDAGLV	1531					
Db	1501	DKGEIO	EGYCAP	SDDLQ	QRG	LFYSMA	KDAGLV	1531					

RESULT 2

Q9U099

ID

Q9U099

PRELIMINARY;

PRT;

1515 AA.

AC

Q9U099;

01-MAY-2000 (TrEMBLrel. 13, Created)

DT

01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT

01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE

Multi drug resistance protein (Fragment).

GN

Name=MRP;

OS

Homo sapiens (Human).

OC

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI\_TaxID=9606;

ON

[1]

RP

SEQUENCE FROM N.A.

RA

MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;

RX

Grant C.E., Kurz E.U., Cole S.P., Deeley R.G.;

RT

"Analysis of the intron-exon organization of the human multidrug-

RT

resistance protein gene (MRP) and alternative splicing of its mRNA.";

RL

Genomics 45:368-378 (1997).

CC

-1- SIMILARITY: Belongs to the ABC transporter family.

DR

EMBL; AF022827; AAB83979.1; JOINED.

DR

EMBL; AF022828; AAB83979.1; JOINED.

DR

EMBL; AF022829; AAB83979.1; JOINED.

DR

EMBL; AF022830; AAB83979.1; JOINED.

DR

EMBL; AF022831; AAB83979.1; JOINED.

DR

EMBL; AF022832; AAB83979.1; JOINED.

DR

EMBL; AF022833; AAB83979.1; JOINED.

DR

EMBL; AF022834; AAB83979.1; JOINED.

DR

EMBL; AF022835; AAB83979.1; JOINED.

DR

EMBL; AF022836; AAB83979.1; JOINED.

DR

EMBL; AF022837; AAB83979.1; JOINED.

DR

EMBL; AF022838; AAB83979.1; JOINED.

DR

EMBL; AF022839; AAB83979.1; JOINED.



```

QY 1517 QRLFYSMADAGLV 1531
DB 1501 QRLFYSMADAGLV 1515

RESULT 3
Q864R9 PRELIMINARY; PRT; 1531 AA.
AC Q864R9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidrug resistance protein 1B.
GN Name=MRP1;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22544876; PubMed=12657726;
RA Codinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RA Dantzig A.H., Perry W.L.;
RT "Cloning and functional characterization of the multidrug resistance-
associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
RL Mol. Cancer Ther. 2:307-316(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Perry W.L. III, Godinot N.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY1456673; AAN65349.1; -.
DR HSP; P08716; 1MT0.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRfams; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN 1.
KW ATP-binding.
SQ SEQUENCE 1531 AA; 171658 MW; 1AE788EFD9EF459 CRC64;

Query Match 77.7%; Score 7695; DB 2; Length 1531;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1493; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

QY 1 MARGFCSADGSDPLMDWNNVTNSNPDKFQNTLVWVPCFYLWACFPFVYLSRH 60
DB 1 MARGFCSADGSDPLMDWNNVTNSNPDKFQNTLVWVPCFYLWACFPFVYLSRH 60
QY 61 DRGYQMTPLNKTALGFLLMVWADLFYFWERSRGIFLAPVLSFTLLGITLLA 120
DB 61 DRGYQMTPLNKTALGFLLMVWADLFYFWERSRGIFLAPVLSFTLLGITLLA 120
QY 121 TFLIQLERRKGVQSSGIMLTFWLVALCALILRSKIMTALKEDVQVDLFRDMTFVYVPS 180

```

QY 1261 VERLKEYSETEKAPWQIOETAPPSPQVGRVFNRYCLRYREDLDFVLRHINTYINGG 1320  
 Db VERLKEYSETEKAPWQIOETAPPSPQVGRVFNRYCLRYREDLDFVLRHINTYINGG 1320  
 QY 1321 EVGVIGVGRGAGKSSLTGLFRINSAEGEIIIDGINIAKIGHDLRFXKTIIPQDPVLF 1380  
 Db EVGVIGVGRGAGKSSLTGLFRINSAEGEIIIDGINIAKIGHDLRFXKTIIPQDPVLF 1380  
 QY 1381 SSSLRNLDLPPFQSDDEEVTWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGQRLVCL 1440  
 Db SSSLRNLDLPPFQSDDEEVTWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGQRLVCL 1440  
 QY 1441 ABALLRKTKILVLDATAVDETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db ABALLRKTKILVLDATAVDETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DRKEIQEYGAPSDLLQQRGLFYFMAKDALV 1531  
 Db DRKEIQEYGAPSDLLQQRGLFYFMAKDALV 1531

## RESULT 4

Q864S0 PRELIMINARY; PRT; 1531 AA.  
 AC Q864S0;  
 DT 01-JUN-2003 (Tremblrel. 24, Created)  
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
 DE Multidrug resistance protein 1A.  
 GN Name=MRP1;  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
 OC Cercopitheciae; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22544876; PubMed=12657726;  
 RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,  
 RA Dantzig A.H., Perry W.L.;  
 RT "Cloning and functional characterization of the multidrug resistance-  
 associated protein (MRP1/ABCC1) from the cynomolgus monkey.";  
 RL Mol. Cancer Ther. 2:307-316(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Perry W.L. III, Godinot N.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AY146672; AAN65348.1; -;  
 DR HSP; P08716; IMT0.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005524; F:ATPase activity; IEA.  
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; F:transport; IEA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR011527; ABC membrane 1.  
 DR InterPro; IPR001140; ABC TM transp.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR005292; MRP assoc.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR Pfam; PF00664; ABC membrane; 2.  
 DR Pfam; PF00005; ABC tran; 2.  
 DR ProDom; PD000006; ABC transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAMs; TIGR00957; MRP assoc\_pro; 1.  
 DR PROSITE; PS50929; ABC TMIF; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 2.

DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 KW ATP-binding.  
 SQ SEQUENCE 1531 AA; 171700 MW; 168712B8CC2D2B89 CRC64;  
 Query Match 77.6%; Score 7689; DB 2; Length 1531;  
 Best Local Similarity 97.6%; Pred. No. 0;  
 Matches 1494; Conservative 18; Mismatches 19; Indels 0; Gaps 0;  
 QY 1 MALRGFCADGSDPLMDMNTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFYLYLSRH 60  
 Db MALRGFCADGSDPLMDMNTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFYLYLSRH 60  
 QY 61 DRGIQMTPIAKTKTALGFLMIWCWADLFYSFWERSRGIFLAPVFLVSPFLGLTLLA 120  
 Db DRGIQMTPLAKTKTALGFLMIWCWADLFYSFWERSRGIFLAPVFLVSPFLGLTLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVLCALAILASKINTALKEDAQVDFLDRIDIFYVYFS 180  
 Db TFLIQLERRKGVQSSGIMLTFWLVLCALAILASKINTALKEDAQVDFLDRIDIFYVYFS 180  
 QY 181 LLLIQLVLSGSDRSPFSETIHDNPNCPSSASFLSRITFWWITGLIVRGYRQPLEGSD 240  
 Db LVLIQLVLSGSDRSPFSETIHDNPNCPSSASFLSRITFWWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSEOVVPLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKYDANEEVEAL 300  
 Db LWSLNKEDTSEOVVPLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKYDANEEVEAL 300  
 QY 301 IVKSPQKEWNPSPFKVLYKTFPGPYFLMSFFPKAIDHLMFSGPQILKLIKFNVDTKAPD 360  
 Db IVKSPQKEWNPSPFKVLYKTFPGPYFLMSFFPKAIDHLMFSGPQILKLIKFNVDTKAPD 360  
 QY 361 WQGYFYVLLFVTVACLOTVLVHQYFHICFVSGMRKTAIVGAVYRKALVITNSARKSTV 420  
 Db WQGYFYTALLFVAACLOTVLVHQYFHICFVSGMRKTAIVGAVYRKALVITNSARKSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLLMLNLGSLVAGVAVMLVMPVN 480  
 Db GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLLMLNLGSLVAGVAVMLVMPVN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYAWELAFKDKVLAIROBELKVLK 540  
 Db AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAVLSAVGFTWVCTPFLVALCTFAVYVITDENIILDAQTAFAVLSALFNLRPLNLTP 600  
 Db KSAVLAAGFTWVCTPFLVALCTFAVYVITDENIILDAQTAFAVLSALFNLRPLNLTP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSITERRPVKGGTNSITVRNATFTWARSDDPT 660  
 Db MVISSIVQASVSLKRLRIFLSHEELEPDSITERRPVKGGTNSITVRNATFTWARSDDPT 660  
 QY 661 LNGITFSIPEGALVAVVGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQI 720  
 Db LNGITFSIPEGALVAVVGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQI 720  
 QY 721 SLRNLFGCOLEBPYRVIQACALLPDLEILPSGDRTEIGKGVNLSGCKOKORVSLAR 780  
 Db SLQENILFGCOLEBPYRVIQACALLPDLEILPSGDRTEIGKGVNLSGCKOKORVSLAR 780  
 QY 781 AVSNADIYLFDDPLSAVDAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840  
 Db AVYCNADIYLFDDPLSAVDAHVGHKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFBLRTYASTEQDAENGVGTGVSQPGKEAKQEMNGM 900  
 Db MSGGKISEMGSYQELLARDGAFBLRTYASTEQDAENGVGTGVSQPGKEAKQEMNGM 900  
 QY 901 LVTDGAKQLOQLSSSSSSSGDLSRHNTAEILOKAEKEETWKLMEADKATQGVK 960  
 Db LVTDGAKQLOQLSSSSSSSGDLSRHNTAEILOKAEKEETWKLMEADKATQGVK 960  
 QY 961 SVYWDYMKATGLFISFLSIPLFMCHVSNALASNYLWLSLWTDVNGTQHTKVRLSVYG 1020

```

Db 961 SVYDYMKAIGLFISFLSIFLFCNHWAAASNYLWLTDDPVGNGTOEHTKVLRSVYG 1020
QY 1021 ALGISGIAVGYMAVSGGILASRLHVDLLHSLRSPMSFFERTPSGNLVNRFPSKEL 1080
Db 1021 ALGISGIAVGYMAVSGGILASRLHVDLLHSLRSPMSFFERTPSGNLVNRFPSKEL 1080
QY 1081 DTVDMSIPEVIMKMGSLFNVCACIVILLATPIAAIIPPLGLIYFFVQRFVYASSRQL 1140
Db 1081 DTVDMSIPEVIMKMGSLFNVCACIVILLATPIAAIIPPLGLIYFFVQRFVYASSRQL 1140
QY 1141 KLESVSRSPVYSHNETLLGVSIVRAPEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200
Db 1141 KLESVSRSPVYSHNETLLGVSIVRAPEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200
QY 1201 VRLCEVGNICIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNLVMSSEMETNIVA 1260
Db 1201 VRLCEVGNICIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEAPMQIQTAPPSPQVGRVFNRYNCLRYREDLDFVLRHINVTINGG 1320
Db 1261 VERLKEYSETEAPMQIQTAPPSPQVGRVFNRYNCLRYREDLDFVLRHINVTINGG 1320
QY 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380
QY 1381 SGLSRNLNLPFFSOYDEEVWTSLELAHLKDFVSALPDKLDHCAEGENLSGQRLVCL 1440
Db 1381 SGLSRNLNLPFFSOYDEEVWTSLELAHLKDFVSALPDKLDHCAEGENLSGQRLVCL 1440
QY 1441 ARALLRKTILVDEATAVLETDLQIQTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 ARALLRKTILVDEATAVLETDLQIQTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DRGETQYEGAPSDLLQOQGLFYSMAKDAGLV 1531
Db 1501 DRGETQYEGAPSDLLQOQGLFYSMAKDAGLV 1531

RESULT 5
Q9UQ97 PRELIMINARY; PRT: 1459 AA.
AC Q9UQ97;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidrug resistance protein (fragment).
GN Name=MRP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.E., Kurz E.U., Cole S.P., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
RT resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378(1997).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF022824; AAB83980.1; JOINED.
DR EMBL; AF022825; AAB83980.1; JOINED.
DR EMBL; AF022826; AAB83980.1; JOINED.
DR EMBL; AF022827; AAB83980.1; JOINED.
DR EMBL; AF022828; AAB83980.1; JOINED.
DR EMBL; AF022829; AAB83980.1; JOINED.
DR EMBL; AF022830; AAB83980.1; JOINED.
DR EMBL; AF022831; AAB83980.1; JOINED.
DR EMBL; AF022832; AAB83980.1; JOINED.
DR EMBL; AF022833; AAB83980.1; JOINED.
DR EMBL; AF022834; AAB83980.1; JOINED.
DR EMBL; AF022835; AAB83980.1; JOINED.

```

```

DR EMBL; AF022836; AAB83980.1; JOINED.
DR EMBL; AF022837; AAB83980.1; JOINED.
DR EMBL; AF022838; AAB83980.1; JOINED.
DR EMBL; AF022839; AAB83980.1; JOINED.
DR EMBL; AF022841; AAB83980.1; JOINED.
DR EMBL; AF022842; AAB83980.1; JOINED.
DR EMBL; AF022843; AAB83980.1; JOINED.
DR EMBL; AF022844; AAB83980.1; JOINED.
DR EMBL; AF022845; AAB83980.1; JOINED.
DR EMBL; AF022846; AAB83980.1; JOINED.
DR EMBL; AF022847; AAB83980.1; JOINED.
DR EMBL; AF022848; AAB83980.1; JOINED.
DR EMBL; AF022849; AAB83980.1; JOINED.
DR EMBL; AF022850; AAB83980.1; JOINED.
DR EMBL; AF022851; AAB83980.1; JOINED.
DR EMBL; AF022852; AAB83980.1; JOINED.
DR EMBL; AF022853; AAB83980.1; JOINED.
DR HSSP; P08716; 1MT0.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0042626; F: ATPase activity, coupled to transmembrane m...; IEA.
DR GO; GO:0000166; F: nucleotide binding; IEA.
DR GO; GO:0005215; F: transporter activity; IEA.
DR GO; GO:0006810; F: transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR011527; ABC membrane 1.
DR InterPro; IPR011140; ABC TM transp.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC membrane; 2.
DR Pfam; PF00005; ABC tran; 1.
DR ProDom; PD000006; ABC transporter; 1.
DR SMART; SM00382; AAA; 2.
DR TIGRfams; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding.
FT NON_TER
SQ SEQUENCE 1459 AA; 163830 MW; 6A016A5A6E7AA95C CRC64;

Query Match 75.2%; Score 7444; DB 2; Length 1459;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1459; Conservative 0; Mismatches 0; Indels 56; Gaps 1;

QY 17 DNNVTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFYLILSRHGRGIQMTPLNKTITA 76
Db 1 DNNVTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFYLILSRHGRGIQMTPLNKTITA 60
QY 77 LGFLLLWIVCWADLFYSFWRSGRIELAPVFLVSPDLLGTTLLATFLQLERRKGVQSSG 136
Db 61 LGFLLLWIVCWADLFYSFWRSGRIELAPVFLVSPDLLGTTLLATFLQLERRKGVQSSG 120
QY 137 IMLTFWLVALCALAILRASKIMTALKEDAQVDLFRDITFYVYFVSLLLQLVLSCFSDRSP 196
Db 121 IMLTFWLVALCALAILRASKIMTALKEDAQVDLFRDITFYVYFVSLLLQLVLSCFSDRSP 180
QY 197 LFSETHDPNCPSSASFLSRITFWITGLIVRGYRQPLEGSDLWSLNKETSQVVPV 256
Db 181 LFSETHDPNCPSSASFLSRITFWITGLIVRGYRQPLEGSDLWSLNKETSQVVPV 240
QY 257 LVKNWKKCAKTRKQPVKVYVSSKDPAPQKSSKVDANEVEALIVKSPQKWNPSLFKV 316
Db 241 LVKNWKKCAKTRKQPVKVYVSSKDPAPQKSSKVDANEVEALIVKSPQKWNPSLFKV 300
QY 317 LYKTGPGYFLMSFFPKAHLDMFSGPQLKLLIKFVNDTKAPDQWQGYFYTVLLFVTACL 376
Db 301 LYKTGPGYFLMSFFPKAHLDMFSGPQLKLLIKFVNDTKAPDQWQGYFYTVLLFVTACL 360
QY 377 QTLVLHQYFHICFVSGMRKITAIVGAVYRKALVIYNSARKSSTVGIEIVNLMVSDAQRWMD 436

```



Db 361 QTLVHQYFHCIFVSGMIRIKTAVIGAVYRKALVITNSARKSVTVGEIVNLSVDAQRFMD 420  
 QY 437 LATYINMIWSAPLOVTLALYLWNLGSPSVLAGVAVMLVMPVNAVAMKTKTYOVAYHK 496  
 Db 421 LATYINMIWSAPLOVTLALYLWNLGSPSVLAGVAVMLVMPVNAVAMKTKTYOVAYHK 480  
 QY 497 SKDNRIKLMNEILNGIKVLKYLAWELAFKDKVLAIROBELKVLKKSAYLSAVGTFTWCT 556  
 Db 481 SKDNRIKLMNEILNGIKVLKYLAWELAFKDKVLAIROBELKVLKKSAYLSAVGTFTWCT 540  
 QY 557 PFLVALCTFAVYVTDENNILDAQAFVSLALFNILRPPNLPMPVSSIVQASVSLKRL 616  
 Db 541 PFLVALCTFAVYVTDENNILDAQAFVSLALFNILRPPNLPMPVSSIVQASVSLKRL 600  
 QY 617 RFLSHEELEPDSIERPPVKDGGTNSITVRNATFTWASDPTPLNGITFTSPEGALVAV 676  
 Db 601 RFLSHEELEPDSIERPPVKDGGTNSITVRNATFTWASDPTPLNGITFTSPEGALVAV 660  
 QY 677 VQVCGKSSLSALLAEKMDKVEGHVAKGSAVYVPPQAWIQNDSIRENIFLFGCOLEBPY 736  
 Db 661 VQVCGKSSLSALLAEKMDKVEGHVAKGSAVYVPPQAWIQNDSIRENIFLFGCOLEBPY 720  
 QY 737 YRSVTOACALLPDLEILPSGDRTEIGEKGVNLSGGQKORVSLARAVYSNADLYLDFDPLS 796  
 Db 721 YRSVTOACALLPDLEILPSGDRTEIGEKGVNLSGGQKORVSLARAVYSNADLYLDFDPLS 747  
 QY 797 AVDAHVGKHI FENVIGPKMGLKNKTRILVTHSMSYLPQVDVLIIVMSGGKISMGVQELL 856  
 Db 748 -----KTRILVTHSMSYLPQVDVLIIVMSGGKISMGVQELL 784  
 QY 857 ARDGAFAEFLRTYASTEQEQDAENGVTGSPGKEAKQEMGLMVLTPDSAGKQLQRLSS 916  
 Db 785 ARDGAFAEFLRTYASTEQEQDAENGVTGSPGKEAKQEMGLMVLTPDSAGKQLQRLSS 844  
 QY 917 SSSYSGDISRHNSAEIKQAKAEETWKLMEADKAQTQVKLSVYWDYKAIGLFTSF 976  
 Db 845 SSSYSGDISRHNSAEIKQAKAEETWKLMEADKAQTQVKLSVYWDYKAIGLFTSF 904  
 QY 977 LSIFLFMCNHSALASNTYLSLWTDPIVNGTQHTKVLRSVYGALGISQGLAVFGYSMA 1036  
 Db 905 LSIFLFMCNHSALASNTYLSLWTDPIVNGTQHTKVLRSVYGALGISQGLAVFGYSMA 964  
 QY 1037 VSIGILASRCLHVDLLSHILRSPMSFERTPTSGNLVNRFSKELTVDMSIPEVIKMPMG 1096  
 Db 965 VSIGILASRCLHVDLLSHILRSPMSFERTPTSGNLVNRFSKELTVDMSIPEVIKMPMG 1024  
 QY 1097 SLFNIVGACIVLLATPTAAIIIPGLIYFPVQRFYVASSRQLKLESVRSRSPYSHFN 1156  
 Db 1025 SLFNIVGACIVLLATPTAAIIIPGLIYFPVQRFYVASSRQLKLESVRSRSPYSHFN 1084  
 QY 1157 ETLGLSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLAVRLCEVNCIVLFAAL 1216  
 Db 1085 ETLGLSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLAVRLCEVNCIVLFAAL 1144  
 QY 1217 FAVISRHSLSAGLVGLSVYSQVTTYLNLVRMSSEMETNIVAVRLKEYSETEKEAPW 1276  
 Db 1145 FAVISRHSLSAGLVGLSVYSQVTTYLNLVRMSSEMETNIVAVRLKEYSETEKEAPW 1204  
 QY 1277 QIQTAPPSWPOVGRVFRNYCLRYREDLPVLRHNTINGGKVIQVGTGAKGSSL 1336  
 Db 1205 QIQTAPPSWPOVGRVFRNYCLRYREDLPVLRHNTINGGKVIQVGTGAKGSSL 1264  
 QY 1337 TLGLFRINESAGEIIDIIGINIAKIGLHDLAFKTIIPQDPVLPFSGSLRMNLDPFQSYSD 1396  
 Db 1265 TLGLFRINESAGEIIDIIGINIAKIGLHDLAFKTIIPQDPVLPFSGSLRMNLDPFQSYSD 1324  
 QY 1397 EEVWTSLELAHLKDPVSLPDKLHDECAEGENLSVGQRLVCLARLLRTKTIILVDEA 1456  
 Db 1325 EEVWTSLELAHLKDPVSLPDKLHDECAEGENLSVGQRLVCLARLLRTKTIILVDEA 1384  
 QY 1457 TAAVLETDLLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVILVDKGEIQVYGAPSDLLQ 1516  
 Db 1385 TAAVLETDLLIQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVILVDKGEIQVYGAPSDLLQ 1444

QY 1517 ORGLFYSNAKDAGLV 1531  
 Db 1445 ORGLFYSNAKDAGLV 1459  
 RESULT 6  
 QSUQA0  
 ID Q9UQA0 PRELIMINARY; PRT; 1456 AA.  
 AC Q9UQA0;  
 DT 01-MAY-2000 (TREMELrel. 13, Created)  
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMELrel. 26, Last annotation update)  
 DE Multidrug resistance protein (Fragment).  
 GN Name=MRP;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;  
 RA Grant C.E., Kurz E.U., Cole S.P., Deeley R.G.;  
 RT "Analysis of the intron-exon organization of the human multidrug-  
 resistance protein gene (MRP) and alternative splicing of its mRNA.";  
 RL Genomics 45:368-378(1997).  
 CC -1- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AF022824; AAB83981.1; JOINED.  
 DR EMBL; AF022825; AAB83981.1; JOINED.  
 DR EMBL; AF022826; AAB83981.1; JOINED.  
 DR EMBL; AF022827; AAB83981.1; JOINED.  
 DR EMBL; AF022828; AAB83981.1; JOINED.  
 DR EMBL; AF022829; AAB83981.1; JOINED.  
 DR EMBL; AF022830; AAB83981.1; JOINED.  
 DR EMBL; AF022831; AAB83981.1; JOINED.  
 DR EMBL; AF022832; AAB83981.1; JOINED.  
 DR EMBL; AF022833; AAB83981.1; JOINED.  
 DR EMBL; AF022834; AAB83981.1; JOINED.  
 DR EMBL; AF022835; AAB83981.1; JOINED.  
 DR EMBL; AF022836; AAB83981.1; JOINED.  
 DR EMBL; AF022837; AAB83981.1; JOINED.  
 DR EMBL; AF022838; AAB83981.1; JOINED.  
 DR EMBL; AF022840; AAB83981.1; JOINED.  
 DR EMBL; AF022841; AAB83981.1; JOINED.  
 DR EMBL; AF022842; AAB83981.1; JOINED.  
 DR EMBL; AF022843; AAB83981.1; JOINED.  
 DR EMBL; AF022844; AAB83981.1; JOINED.  
 DR EMBL; AF022845; AAB83981.1; JOINED.  
 DR EMBL; AF022846; AAB83981.1; JOINED.  
 DR EMBL; AF022847; AAB83981.1; JOINED.  
 DR EMBL; AF022848; AAB83981.1; JOINED.  
 DR EMBL; AF022849; AAB83981.1; JOINED.  
 DR EMBL; AF022850; AAB83981.1; JOINED.  
 DR EMBL; AF022851; AAB83981.1; JOINED.  
 DR EMBL; AF022852; AAB83981.1; JOINED.  
 DR EMBL; AF022853; AAB83981.1; JOINED.  
 DR HSP; P08716; 1MT0.  
 DR GO; GO:0016021; C: integral to membrane; IEA.  
 DR GO; GO:0005524; F: ATP binding; IEA.  
 DR GO; GO:0042626; F: ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO; GO:0000166; F: nucleotide binding; IEA.  
 DR GO; GO:0005215; F: transporter activity; IEA.  
 DR GO; GO:0006810; P: transport; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR011527; ABC\_membrane\_1.  
 DR InterPro; IPR001140; ABC\_TM\_transp.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR Pfam; PF00664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transporter; 2.  
 DR SMART; SM00382; AAA; 2.



DR TIGRAME; TIGR00957; MRP\_abcoc\_pro; 1.  
 DR PROSITE; P50929; ABC\_TMIF; 2.  
 DR PROSITE; P50211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; P50893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE; P500107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 KW ATP-binding.  
 FT NON\_TER  
 SQ SEQUENCE 1456 AA; 163231 MW; 8DE8AAB22BC481P2 CRC64;  
 Query Match 74.9%; Score 7419.5; DB 2; Length 1456;  
 Best Local Similarity 96.1%; Pred. No. 0;  
 Matches 1456; Conservative 0; Mismatches 0; Indels 59; Gaps 1;  
 QY 17 DNNVTWNTSNDPFCFQNTLVWVPCFYLWACFFPYFLYLSRHRDGYQMTPLNKTCTA 76  
 DB 1 DNNVTWNTSNDPFCFQNTLVWVPCFYLWACFFPYFLYLSRHRDGYQMTPLNKTCTA 60  
 QY 77 LGFLWIVCWADLFYSFWEERSRGIFLAPVFLVSPILLGITTLATFLQLERRKGVQSSG 136  
 DB 61 LGFLWIVCWADLFYSFWEERSRGIFLAPVFLVSPILLGITTLATFLQLERRKGVQSSG 120  
 QY 137 IMLTWLVVALVCAALILRSKIMTALKEDAQVDFRDIIFYVYFSLILLQLVLSCFSDRSP 196  
 DB 121 IMLTWLVVALVCAALILRSKIMTALKEDAQVDFRDIIFYVYFSLILLQLVLSCFSDRSP 180  
 QY 197 LFSETHDPNCPCESSASPLSRITFWITGLIVRGYQPLEGSDLSLWLNKEDTSEQVVPV 256  
 DB 181 LFSETHDPNCPCESSASPLSRITFWITGLIVRGYQPLEGSDLSLWLNKEDTSEQVVPV 240  
 QY 257 LVNKKKECAKTRKOPKPVVYSSKDPAPQKESKVDANEVEALIVKSPQKWNPSLFKV 316  
 DB 241 LVNKKKECAKTRKOPKPVVYSSKDPAPQKESKVDANEVEALIVKSPQKWNPSLFKV 300  
 QY 317 LYKTFGPFLMFFPKAHLDMFSGPQLKLLIKFVNDTKAPDMQGYFTYVLLFVTAQL 376  
 DB 301 LYKTFGPFLMFFPKAHLDMFSGPQLKLLIKFVNDTKAPDMQGYFTYVLLFVTAQL 360  
 QY 377 QTLVLHQPHICFVSGMRITKAVIGAVYRKALVITNSARKSSTVGEIVNLSVDAQRPM 436  
 DB 361 QTLVLHQPHICFVSGMRITKAVIGAVYRKALVITNSARKSSTVGEIVNLSVDAQRPM 420  
 QY 437 LATYINMISAPLOVITLALYLWNLGPSVLGAVVMVLPVNAVMAKTKTYOVAHMK 496  
 DB 421 LATYINMISAPLOVITLALYLWNLGPSVLGAVVMVLPVNAVMAKTKTYOVAHMK 480  
 QY 497 SKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIRQBELKYLKSAVLSAVGTFTWCT 556  
 DB 481 SKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIRQBELKYLKSAVLSAVGTFTWCT 540  
 QY 557 PELVALCTPAVVTIDENNILDAQAFVSLALFNILRFPNLPWVISSIVQASVSLKRL 616  
 DB 541 PELVALCTPAVVTIDENNILDAQAFVSLALFNILRFPNLPWVISSIVQASVSLKRL 600  
 QY 617 RIFLSHEELEPDSIBERRPVKGGTNSITVRNATFTWARSDDPTLNGITFSIPEGALVAV 676  
 DB 601 RIFLSHEELEPDSIBERRPVKGGTNSITVRNATFTWARSDDPTLNGITFSIPEGALVAV 660  
 QY 677 VQGVCGGKSSLSALLAEMDKVEGHVIAIKGSVAYVPOQAWIQNDSIRENILEGCGLEBPY 736  
 DB 661 VQGVCGGKSSLSALLAEMDKVEGHVIAIKGSVAYVPOQAWIQNDSIRENILEGCGLEBPY 688  
 QY 737 YRSVIQACALLPDLITLPSGDRTEIGKGNLSGGOKORVSLARAVYSNADYLFDDPLS 796  
 DB 689 -----KGNLSGGOKORVSLARAVYSNADYLFDDPLS 721  
 QY 797 AVDAHVGKHFENVIPKGMKNKTRILVTHSMYLPQVDVITVMSGGKISMGVQELL 856  
 DB 722 AVDAHVGKHFENVIPKGMKNKTRILVTHSMYLPQVDVITVMSGGKISMGVQELL 781  
 QY 857 ARDGAFAELRYASTEQODAEENGVTGSGPGKEAKOMENGLVTDASAGLQRLSS 916  
 DB 782 ARDGAFAELRYASTEQODAEENGVTGSGPGKEAKOMENGLVTDASAGLQRLSS 841

QY 917 SSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKAOTGVKLSVYWDYMKAIGLFISF 976  
 DB 842 SSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKAOTGVKLSVYWDYMKAIGLFISF 901  
 QY 977 LSIPLFMCNHYVALSNYLSLWTDPIVNGTQOEHKTVRLSVYGALGISQGIATVFGYMA 1036  
 DB 902 LSIPLFMCNHYVALSNYLSLWTDPIVNGTQOEHKTVRLSVYGALGISQGIATVFGYMA 961  
 QY 1037 VSIIGILASRCLHVDLLSHLSILRSPMSFFERTPSGNLVRNFRSKELDTVDSDIMEVIMKFMG 1096  
 DB 962 VSIIGILASRCLHVDLLSHLSILRSPMSFFERTPSGNLVRNFRSKELDTVDSDIMEVIMKFMG 1021  
 QY 1097 SLFNIVGACIVILLATPIAAIIIPPLGLIYFVQRFYVASSRQLKRLSVSRSPYSHFN 1156  
 DB 1022 SLFNIVGACIVILLATPIAAIIIPPLGLIYFVQRFYVASSRQLKRLSVSRSPYSHFN 1081  
 QY 1157 ETLGVSIVIRAFEBQERFIHOSDLKVDENKAYYPSIVANRWLAVRLECVGNCIVLFAAL 1216  
 DB 1082 ETLGVSIVIRAFEBQERFIHOSDLKVDENKAYYPSIVANRWLAVRLECVGNCIVLFAAL 1141  
 QY 1217 FAVISRSHLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVAVERLKEYSETEKEAPW 1276  
 DB 1142 FAVISRSHLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVAVERLKEYSETEKEAPW 1201  
 QY 1277 QIQTAPSSWPQGRVFRNYCLRYREDLDFVLHINVTINGGKVGIVGRTGAGKSSL 1336  
 DB 1202 QIQTAPSSWPQGRVFRNYCLRYREDLDFVLHINVTINGGKVGIVGRTGAGKSSL 1261  
 QY 1337 TLGLFRINESAGEITIIDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMNLDPFSSQYSD 1396  
 DB 1262 TLGLFRINESAGEITIIDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMNLDPFSSQYSD 1321  
 QY 1397 BEVWTSLSLAHLKDFVSALPDKLDECAEGENLSVGQRLVCLARALLRKTILVLDEA 1456  
 DB 1322 BEVWTSLSLAHLKDFVSALPDKLDECAEGENLSVGQRLVCLARALLRKTILVLDEA 1381  
 QY 1457 TAAVLETDLLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVILVDKGEIQEYGAESDILQ 1516  
 DB 1382 TAAVLETDLLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVILVDKGEIQEYGAESDILQ 1441  
 QY 1517 QRGIFYSMKADAGLV 1531  
 DB 1442 QRGIFYSMKADAGLV 1456  
 RESULT 7  
 QUR05  
 ID Q6UR05 PRELIMINARY; PRT; 1531 AA.  
 AC Q6UR05;  
 DT 05-JUL-2004 (TREMELrel. 27, Created)  
 DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)  
 DE Multidrug resistance-associated protein 1.  
 GN Name=MRP1;  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OC NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22404678; PubMed=13516967;  
 RA Ma L., Pratt S.E., Cao J., Dantzig A.H., Moore R.E., Slapak C.A.;  
 RT "Identification and characterization of the canine multidrug  
 resistance-associated protein";  
 RL Mol. Cancer Ther. 1:1335-1342(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ma L., Pratt S.E., Cao J., Dantzig A.H., Moore R.E., Slapak C.A.;  
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AY363728; AAQ3148.1; --  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m...; IEA.  
 DR GO; GO:000166; F:nucleotide binding; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0008610; F:transport; IEA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR011527; ABC membrane 1.  
 DR InterPro; IPR001140; ABC TM transp.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR001395; ABC/ket\_red.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR Pfam; PF00664; ABC membrane; 2.  
 DR Pfam; PF00005; ABC tran; 2.  
 DR ProDom; PD000006; ABC transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRfam; TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE; PS00929; ABC TMIF; 2.  
 DR PROSITE; PS00211; ABC TRANSPORTER 1; 2.  
 DR PROSITE; PS00893; ABC TRANSPORTER 2; 2.  
 DR PROSITE; PS00063; ALDOKETO REDUCTASE 3; UNKNOWN 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN 1.  
 DR ATP-binding.  
 DR SEQUENCE 1531 AA; 171790 MW; 746361A71C6158BD CRC64;  
 SQ  
 Query Match 73.4%; Score 7272; DB 2; Length 1531;  
 Best Local Similarity 91.6%; Pred. No. 0;  
 Matches 1402; Conservative 70; Mismatches 59; Indels 0; Gaps 0;  
 QY 1 MALRGCADGSDPLDWNVNTWNTNPDFTKCFONTVLVWPCFYLMACFPFYLYLSRH 60  
 DB 1 MALRGCADGSDPFWEDVSNWNTNPDFTKCFONTVLVWPCFYLYLSRH 60  
 QY 61 DRGYTQMTPLNKTALGFLWIVCWADLFYSFWERSGILFAPVFLPSLTLLGTTLLA 120  
 DB 61 DRGYTQMTPLNKTALGFLWIVCWADLFYSFWERSGILFAPVFLPSLTLLGTTLLA 120  
 QY 121 TFLQLERRKGQSSGIMLFWALVACALALIRSKIMTALKEDAQVDFRDITPVYVFS 180  
 DB 121 TFLQLERRKGQSSGIMLFWALVACALALIRSKIMTALKEDAEIDVFRDITPVYVFS 180  
 QY 181 LLLQLVLSRSDRPLSETHDNPCESSASFLSRITFWITGLIVRGVROPLEGSD 240  
 DB 181 LVLIQLVLSRSDRPLSETHDNPCESSASFLSRITFWITGLIVRGVROPLEGSD 240  
 QY 241 LWSLNKEDTSEOVVPLVKNWKECAKTRKOPVYVSKDPAQPKESKVDANEEVAL 300  
 DB 241 LWSLNKEDTSEOVVPLVKNWKECAKTRKOPVYVSKDPAQPKESKVDANEEVAL 300  
 QY 301 IVKSPQKEMNPFLFKVLYKTGYPFLMSFFFKAIHDLMMFSGPQLKLLIKFVNDTKAPD 360  
 DB 301 IVKTPKEREPSLFLVLYKTGYPFLMSFFFKAIHDLMMFAGPEILKLLINFVNDKKAPD 360  
 QY 361 WQGYFYTVLLFTACLOTLLVHOYPHICFVSGMRJKTAVIGAVYRKALVITNSARKSTV 420  
 DB 361 WQGYLYTALLFTACLOTLLVHOYPHICFVSGMRJKTAVIGAVYRKALVITNSARKSTV 420  
 QY 421 GEIVNLSVDAORFMDLTYINMWSAPLOVTLALYLWNLGSPVLGAVVWMLVWPN 480  
 DB 421 GEIVNLSVDAORFMDLTYINMWSAPLOVTLALYLWNLGSPVLGAVVWMLVWPN 480  
 QY 481 AVAMAKTKTYQVAHMKSKDNRIKLMEILNGIKVLYAWELAPKDKVLAIROBELKVLK 540  
 DB 481 AVAMAKTKTYQVAHMKSKDNRIKLMEILNGIKVLYAWELAPKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPPLVALCTFAVVTVDENNILDAQTAQAFVSLALFNILRPLNLP 600  
 DB 541 KSAYLSAVGTFTWCTPPLVALCTFAVVTVDENNILDAQTAQAFVSLALFNILRPLNLP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERPPVKDGGGNTSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERPPVKDGGGNTSITVRNATFTWARSDDPT 660  
 QY 661 LNTGTFSPGALVAVGVGCGKSSLLSALLAEMDKVEGHVAKGSVAYVPPQQAWIQND 720

DB 661 LNTGTFSPGALVAVGVGCGKSSLLSALLAEMDKVEGHVAKGSVAYVPPQQAWIQND 720  
 QY 721 SLRENILFGCOLLEBPYRSVIOACALLPDLEILPSGDRTEIGEGWNLGGQKQSVLSAR 780  
 DB 721 SLRENILFGCOLLEBPYRSVIOACALLPDLEILPSGDRTEIGEGWNLGGQKQSVLSAR 780  
 QY 781 AVYNADIIYLFDDPLSAVDHVGKHI FENVIGPKGNLKNKTRILVTHSMVSLPOVDVILV 840  
 DB 781 AVYCDSDIYLFDDPLSAVDHVGKHI FENVIGPKGNLKNKTRILVTHSMVSLPOVDVILV 840  
 QY 841 MSGGKISEMGYSQELLARDGAPAEFLRTYASTEOEDAEENGVTGVSFGPGKEAKOMNGM 900  
 DB 841 MTGGKISEMGYSQELLARDGAPAEFLRTYASGDQEAQDDGLTGVSFGPGKEAKOMNGM 900  
 QY 901 LVTDAGKQLQRLSSSSSYSGDISRHHNSTAEILOKABAKKEETWKLMEADKAQGTQVKL 960  
 DB 901 LVTDVAGKQLQRLSSSSSYSGDISRHHNSTAEILOKABAKKEETWKLMEADKAQGTQVKL 960  
 QY 961 SVYDYMKAIGLFTISFLIFLPMCNHYSALASNYWLSLWTDPIVNGTQEHKTVLSVYG 1020  
 DB 961 SVYDYMKAIGLFTISFLIFLPMCNHYSALASNYWLSLWTDPIVNGTQEHKTVLSVYG 1020  
 QY 1021 ALGISQGIADVFGYSMAVSIIGGILASRCILHVDLLHSILRSPMSFFERTPSGNLVNRFPSKEL 1080  
 DB 1021 ALGISQGITVFGYSMAVSIIGGIFASRRLLHVDLLQNLVSPMSFFERTPSGNLVNRFPSKEL 1080  
 QY 1081 DTVDSMPEVITKPMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQFYVASSRQL 1140  
 DB 1081 DTVDSMIPQVITKPMGSLFNIVGACIIILLATPIASIIIPPLGLIYFFVQFYVASSRQL 1140  
 QY 1141 KRLESVSRSVYSHFNETLLGVSVIRAFEEQERFTHOSDLKVDENOKAYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSVYSHFNETLLGVSVIRAFEEQERFTHOSDLKVDENOKAYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVGSLVSYSLQVTTYLNWLRMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVGSLVSYSLQVTTYLNWLRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKEAPWQIETAPPSSWPQVGRVFRNYCLARYEDLDLPVLRHINTVINGG 1320  
 DB 1261 VERLKEYSETEKEAPWQIETAPPSSWPQVGRVFRNYCLARYEDLDLPVLRHINTVINGG 1320  
 QY 1321 BKVGIVGRTGAGKSLTGLFRINESAGEIIIDGINAKIGLHDLRKTITIPQDPVLF 1380  
 DB 1321 BKVGIVGRTGAGKSLTGLFRINESAGEIIIDGINAKIGLHDLRKTITIPQDPVLF 1380  
 QY 1381 SGSLRMLNLPFSQYSDEEVMTSLELAHLKDFVSALPKDLDECAEGGNSLVGQRLVCL 1440  
 DB 1381 SGSLRMLNLPFSQYSDEEVMTSLELAHLKDFVSALPKDLDECAEGGNSLVGQRLVCL 1440  
 QY 1441 ARALLRKTILVLEATAAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
 DB 1441 ARALLRKTILVLEATAAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
 QY 1501 DKGIIQYVGPASDLQOGLFYSMAKDAGLV 1531  
 DB 1501 DKGIIQYVGPASDLQOGLFYSMAKDAGLV 1531  
 RESULT 8  
 Q8HX05  
 ID Q8HX05 PRELIMINARY; PRT; 1530 AA.  
 AC Q8HX05;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Multidrug resistance protein 1.  
 GN Name=MRP1;  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.

NCBI\_TaxID=9913;  
 [1]  
 SEQUENCE FROM N.A.  
 MEDLINE=22062955; PubMed=12067707; DOI=10.1016/S0014-5793(02)02816-8;  
 Taguchi Y., Sasaki K., Komano T.;  
 RA "functional analysis of MRPI cloned from bovine."  
 RT FEBS Lett. 521:211-213(2002).  
 CC -!- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AB082124; BAC15550.1; -;  
 DR HSP; P08716; LMT0.  
 DR GO; GO:0016021; C: integral to membrane; IEA.  
 DR GO; GO:0005524; F: ATP binding; IEA.  
 DR GO; GO:0043262; F: ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO; GO:0000166; F: nucleotide binding; IEA.  
 DR GO; GO:0005215; F: transporter activity; IEA.  
 DR GO; GO:0006810; P: transport; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR011527; ABC membrane 1.  
 DR InterPro; IPR001140; ABC TM transport.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR Pfam; PF00664; ABC membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAMs; TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE; PS0929; ABC\_TMIF; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 KW ATP-binding.  
 SQ SEQUENCE 1530 AA; 171664 MW; AAE4F92ED7832703 CRC64;  
  
 Query Match 73.0%; Score 7230.5; DB 2; Length 1530;  
 Best Local Similarity 90.7%; Pred No. 0;  
 Matches 1388; Conservative 83; Mismatches 59; Indels 1; Gaps 1;  
  
 QY 1 MALRGFCSADGSDPLMDNMVNTWNTSNPDKFQNTLVVWPCFYLVWACFPFFYLYLSRH 60  
 DB 1 MALRDFCSVDGSDLFWEVNVNTWNTSNPDKFQNTLVVWPCFYLVWACFPFFYLYLSRH 60  
 QY 61 DRGYQMTPLNKTATGALGFLMIVCWADLFYFWERSRGIFLAPVFLVSPITLLGTTLLA 120  
 DB 61 DRGYQMTPLNKTATGALGFLMIVCWADLFYFWERSMGKLLAPVFLVSPITLLGTTLLA 120  
 QY 121 TELIOLERRKGVSQSGIMTLFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TELIQERRRGVSQSGIMTLFWLVALCALAILRSKIMTALKEDARVDVFRDITFYVYFS 180  
 QY 181 LLLIQLVLCFSDRSPFLFSETIHDNPPCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LVLIQLVLCFSDRSPFLFSETINDNPPCPSSASFLSRITFWITGMVQYRQPLESTD 240  
 QY 241 LWSLNKEDTSEQVVPVVLVQWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL 300  
 DB 241 LWSLNKEDTSEQVVPVVLVQWKECAKSKQPVVYSSKDPAPKESKVDNVEAEAL 300  
 QY 301 IVKSPQKWNPSLFVLYKTFPGYFLMSPFFKAIDHLMFSGPOILKLLIKFVNDTKAPD 360  
 DB 301 IVKQKQERDPSLFVLYKTFPGYFLMSPFLFAVHDLMMFAGEPEILKLLINFVNDKKAPE 360  
 QY 361 WQGYFYTVLLFVTAQLTLVHLQYFHICFVSGMRKTAIVGAVRKALVITNSARKSTV 420  
 DB 361 WQGYFYTVLLFVTAQLTLVHLQYFHICFVSGMRKTAIVGAVRKALVITNAARKSTV 420  
 QY 421 GEIVNLSVDAQRFDLATYINNIWSAPLOVIALYLLWNLGSPVLAGVAVMLVMPVN 480  
 DB 421 GEIVNLSVDAQRFDLATYINNIWSAPLOVIALYLLWNLGSPVLAGVAVMLVMPVN 480  
 QY 481 AVAMKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKYAWELAFKDKVLAIQBELKVLK 540  
 DB 481 AVAMKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKYAWELAFKDKVLAIQBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPFLNLP 600

DB 541 KSAYLAAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPFLNLP 600  
 QY 601 MVISSIVQASVSLKELRIFLSHEELEPDSIRRRPVKDGSGTNSITVRNATFTWASDDPT 660  
 DB 601 MVISSIVQASVSLKELRIFLSHEELEPDSIRRRPVKDGSGTNSITVRNATFTWASDDPT 660  
 QY 661 LMGITFISPEGALVAVVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720  
 DB 661 LMGITFISPEGALVAVVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND 720  
 QY 721 SLRENILFGCOLBEPYRSVIOACALLPDLEILPSGDRTEIGEGVNLSSGGQKQVSLAR 780  
 DB 721 SLRENILFGCOLBEPYRSVIOACALLPDLEILPSGDRTEIGEGVNLSSGGQKQVSLAR 780  
 QY 781 AVYSNADILYLPDDPLSAVDHVGKHIFENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIV 840  
 DB 781 AVYCDSDVILLDDPLSAVDHVGKHIFENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIV 840  
 QY 841 MSGGKISEMGYSQELLARDGAFAPLRTYASTEQDQAEENGVTGVSQPGKEAKOMENG 900  
 DB 841 MSGGKISEMGYSQELLARDGAFAPLRTYASTEQDQAEENGVTGVSQPGKEAKOMENG 900  
 QY 901 LVTSAGKQLQRLSSSSSSYSYSGDTSRHNSTAELOKAEAKKEETWKLMEADKAQTQVKL 960  
 DB 901 LVTDTAGKMQRLSSSSSSYSYSGDTSRHNSTAELOKAEAKKEETWKLMEADKAQTQVKL 960  
 QY 961 SVYDYMKAIGLFTSFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 DB 961 SVYDYMKAIGLFTSFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 QY 1021 ALGISQIAGVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTSGNLVNRFSKEL 1080  
 DB 1021 ALGISQIAGVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTSGNLVNRFSKEL 1080  
 QY 1081 DTVDSMIPVIMPMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRPYVASSROL 1140  
 DB 1081 DTVDSMIPVIMPMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRPYVASSROL 1140  
 QY 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVIRSRHSLSAGLVLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVIRSRHSLSAGLVLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKEAPWQIQTAPPSSWPQVGRVFNRYCLYREDLDLDFVLRHINTINGG 1320  
 DB 1261 VERLKEYSETEKEAPWQIQTAPPSSWPQVGRVFNRYCLYREDLDLDFVLRHINTINGG 1320  
 QY 1321 EKVGVGTGTGAGKSLTGLFRINESARGEIIIDGINIAKIGLHDLRFKTIIPQDPVLP 1380  
 DB 1321 EKVGVGTGTGAGKSLTGLFRINESARGEIIIDGINIAKIGLHDLRFKTIIPQDPVLP 1380  
 QY 1381 SGSLRMNLDPSQYSDDEEVTSLAHLKDFVSALPDKLDHECAEGGENLSVGQRLVCL 1440  
 DB 1381 SGSLRMNLDPSQYSDDEEVTSLAHLKDFVSALPDKLDHECAEGGENLSVGQRLVCL 1440  
 QY 1441 ARALLRKTILVLDATAVLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ARALLRKTILVLDATAVLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGETOYEGAPSDLLQQRGLFYSMADKGLV 1531  
 DB 1501 DKGETOYEGAPSDLLQQRGLFYSMADKGLV 1531

RESULT 9  
 Q9UQ98  
 ID Q9UQ98 PRELIMINARY; PRT; 1400 AA.  
 AC Q9UQ98;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)



QY 1157 ETLGVSIVRAPEQERFIHQSDLKVDENOKAYPSIVANRWLVRLECVGNCIVLFAAL 1216  
 |||||  
 DB 1026 ETLGVSIVRAPEQERFIHQSDLKVDENOKAYPSIVANRWLVRLECVGNCIVLFAAL 1085  
 |||||  
 QY 1217 FAVISRHSLSAGLVGSYSQVTTYNLWVRMSEMETNIVAVERLKEYSETEKEAPW 1276  
 |||||  
 DB 1086 FAVISRHSLSAGLVGSYSQVTTYNLWVRMSEMETNIVAVERLKEYSETEKEAPW 1145  
 |||||  
 QY 1277 QIQETAPPSWQGVQVEPRNYCLRYREDLDFVLRIHNTVINGEKVGIVGRTGAKGSSL 1336  
 |||||  
 DB 1146 QIQETAPPSWQGVQVEPRNYCLRYREDLDFVLRIHNTVINGEKVGIVGRTGAKGSSL 1205  
 |||||  
 QY 1337 TLGLFRINSAGEIIIDGINIAKIGLHDLRFKTIIPDPVLFSGSLRMNLDPPFSQYSD 1396  
 |||||  
 DB 1206 TLGLFRINSAGEIIIDGINIAKIGLHDLRFKTIIPDPVLFSGSLRMNLDPPFSQYSD 1265  
 |||||  
 QY 1397 EEWTSLELAHLKDFVSALPKDLHCEAGGENSELGVQRLVCLARALLKTKILVLDEA 1456  
 |||||  
 DB 1266 EEWTSLELAHLKDFVSALPKDLHCEAGGENSELGVQRLVCLARALLKTKILVLDEA 1325  
 |||||  
 QY 1457 TAAVLETDLLQSIQRTQFECTVLTIAHRLNTIMDYTRVILDKGEIQEYGAPESDLLQ 1516  
 |||||  
 DB 1326 TAAVLETDLLQSIQRTQFECTVLTIAHRLNTIMDYTRVILDKGEIQEYGAPESDLLQ 1385  
 |||||  
 QY 1517 QRLGYFSMAKADGLV 1531  
 |||||  
 DB 1386 QRLGYFSMAKADGLV 1400  
 |||||

RESULT 10  
 MRPI\_MOUSE  
 ID MRPI\_MOUSE STANDARD; PRT; 1528 AA.  
 AC Q35379;  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Multidrug resistance-associated protein 1 (ATP-binding cassette, sub-family C, member 1).  
 GN Name=Abcc1; Synonyms=Abcc1a, Abcc1b, Mdrap, Mip;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;  
 RX MEDLINE=96251691; PubMed=8649356;  
 RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P., Deeley R.G.;  
 RT "Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter.";  
 RL Mol. Pharmacol. 49:962-971(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S., Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L., Koyagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Nunata K., Okido T., Pavan W.J., Perteira G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M., Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,

RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., Yabuuchi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;  
 RA "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";  
 RT Nature 420:563-573(2002).  
 RL Nature 420:563-573(2002).  
 CC -!- FUNCTION: May participate directly in the active transport of drugs into subcellular organelles or influence drug distribution indirectly (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR EMBL; AF022908; AAB80938.1; -;  
 DR EMBL; AK029876; BAC26654.1; -;  
 DR HSP; F08716; IMT0.  
 DR MGD; MGI:102676; Abcc1.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR011527; ABC membrane 1.  
 DR InterPro; IPR011140; ABC\_TM\_transp.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR Pfam; PF00664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAMs; TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE; PS5029; ABC\_TMIF; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 KW ATP-binding; Glycoprotein; Repeat; Transmembrane; Transport.  
 FT DOMAIN 1 33 Extracellular (By similarity).  
 FT TRANSMEM 34 54 1 (By similarity).  
 FT DOMAIN 55 74 Cytoplasmic (By similarity).  
 FT TRANSMEM 75 95 2 (By similarity).  
 FT DOMAIN 96 100 Extracellular (By similarity).  
 FT TRANSMEM 101 121 3 (By similarity).  
 FT DOMAIN 122 133 Cytoplasmic (By similarity).  
 FT TRANSMEM 134 154 4 (By similarity).  
 FT DOMAIN 155 172 Extracellular (By similarity).  
 FT TRANSMEM 173 193 5 (By similarity).  
 FT DOMAIN 194 317 Cytoplasmic (By similarity).  
 FT TRANSMEM 318 338 6 (By similarity).  
 FT DOMAIN 339 364 Extracellular (By similarity).  
 FT TRANSMEM 365 385 7 (By similarity).  
 FT DOMAIN 386 441 Cytoplasmic (By similarity).  
 FT TRANSMEM 442 462 8 (By similarity).  
 FT DOMAIN 463 465 Extracellular (By similarity).  
 FT TRANSMEM 466 486 9 (By similarity).  
 FT DOMAIN 487 548 Cytoplasmic (By similarity).  
 FT TRANSMEM 549 569 10 (By similarity).  
 FT DOMAIN 570 591 Extracellular (By similarity).  
 FT TRANSMEM 592 612 11 (By similarity).  
 FT DOMAIN 613 963 Cytoplasmic (By similarity).  
 FT TRANSMEM 964 984 12 (By similarity).  
 FT DOMAIN 985 1022 Extracellular (By similarity).  
 FT TRANSMEM 1023 1043 13 (By similarity).  
 FT DOMAIN 1044 1086 Cytoplasmic (By similarity).  
 FT TRANSMEM 1087 1107 14 (By similarity).

RESULT 11

DR GO: 0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO: 0000166; F:nucleotide binding; IEA.  
 DR GO: 0005215; P:transporter activity; IEA.  
 DR GO: 0006810; P:transport; IEA.  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR011527; ABC\_membrane\_1.  
 DR InterPro: IPR001140; ABC\_TM\_transp.  
 DR InterPro: IPR001439; ABC\_transporter.  
 DR InterPro: IPR001395; Aldol/ket red.  
 DR InterPro: IPR005292; MRP\_assoc.  
 DR Pfam: PF00664; ABC\_membrane; 2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR ProDom: PD00006; ABC\_transporter; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR TIGRFAMs: TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE: PS00211; ABC\_TMIF; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE: PS00893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE: PS00893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE: PS00893; ABC\_TRANSPORTER\_2; 2.  
 DR ATP-binding.  
 KW ATP-binding.  
 SQ SEQUENCE 1532 AA; 171491 MW; 2E6939F63F5A3F68 CRC64;

Query Match 70.0%; Score 6932.5; DB 2; Length 1532;  
 Best Local Similarity 87.3%; Pred. No. 0;  
 Matches 1339; Conservative 99; Mismatches 92; Indels 3; Gaps 3;

QY 1 MALRGCAGSDGDPDLMDNVTWNTGNPDTKCFQNTVLVWVPCFYLWACFPYFLYLSRH 60  
 DB 1 MALSFCSGSDGDPDLMDNVTWNTGNPDTKCFQNTVLVWVPCFYLWACFPYFLYLSRH 60

QY 61 DRGYTOMTLNKTALGFLMIVCWADLFYSFWERSRGIPLAFVFLVSPITLLGTTLLA 120  
 DB 61 DRGYTOMTLNKTALGFLMIVCWADLFYSFWERSRGIPLAFVFLVSPITLLGTTLLA 120

QY 121 TELIQLERRKGVOSSGIMLTFWLAVCALILRSKIMTALKEDAQVDLPDITFYVYPS 180  
 DB 121 TELIQLERRKGVOSSGIMLTFWLAVCALILRSKIMTALKEDAQVDLPDITFYVYPS 180

QY 181 LILLIQLVLSCFSDRPLFSETTHDPNCPPESSASFLSRITFWITGLIVRGYRQLEGSD 240  
 DB 181 LVFIQLVLSCFSDSPLSETVRDNPDPPESSASFLSRITFWITGLIVRGYRQLEGSD 240

QY 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPVKVVYS - SKDPAQPKESSKVDANEVEA 299  
 DB 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPVKVVYS - SKDPAQPKESSKVDANEVEA 299

QY 300 LIVKSPQKWNPSLKVLYKTFPGPYFLMSFPFKAIHDLMMSPGPOILKLLIKFVNDTKAP 359  
 DB 301 LIVKSKHNDPDLSPFKVLYKTFPGPYFLMSFLYKALHDLMMFAGPBEILELIINFVNDREAP 360

QY 360 DWQGYFYTVLFTVLTACLOTVLHQLYFHICFVSGMRIKTAIVIGAVYRKALVITNSARKSST 419  
 DB 361 DWQGYLYTALLFVSLACLOTLALHQLYFHICFVSGMRIKTAIVIGAVYRKALVITNSARKSST 420

QY 420 VGEIVNLMSVDQAFMDLATYINMTWSAPLQVILALYLNLNGLSPSLAGVAVMLVMPV 479  
 DB 421 VGEIVNLMSVDQAFMDLATYINMTWSAPLQVILALYLNLNGLSPSLAGVAVMLVMPV 480

QY 480 NAWMAKTKTYQVAHMKSKDNRIKLWNLINGIKVLKYLAWEALAPKDKVLAIRQBELKVL 539  
 DB 481 NAWMAKTKTYQVAHMKSKDNRIKLWNLINGIKVLKYLAWEALAPKDKVLAIRQBELKVL 540

QY 540 KKSAYLSAVGTFTWCTPPLVALCTFAVVVTIDENNLDAOTAFVSLALFNILRPLNIL 599  
 DB 541 KKSAYLAAGTFTWCTPPLVALSTFAVVTVDENNLDAKAFVSLALFNILRPLNIL 600

QY 600 PMVISSIVQASVSLKRLRIFLSHEELEPDSTERRPVDGGGTNSITVRNATFTWARSDDP 659  
 DB 601 PMVISSIVQASVSLKRLRIFLSHEELEPDSTERRPVDGGGTNSITVRNATFTWARSDDP 660

QY 660 TLNGITFSPGALVAVGVQCGCKSSLLSALLAEMDKVEGHVATKGSVAVYVQQAQWQ 719  
 DB 661 TLNGITFPAIDGALVAVGVQCGCKSSLLSALLAEMDKVEGHVATKGSVAVYVQQAQWQ 720

QY 720 DSLRENILFGQQLBEPYRYSVIAQCALIPDLLEILPSGDRTEIGKGVNLGGQKQVSLA 779  
 DB 721 DSLRENILFGRLQEHCHYKAVMEACALLPDLLEILPSGDLTEIGKGVNLGGQKQVSLA 780

QY 780 RAVYSNADIYLPDPLSAVDHVGKHIPENVIGPKGMLKNKTRILVITHSMVSLPOVDVII 839  
 DB 781 RAVYCNSDIYLLDDPLSAVDHVGKHIPENVIGPKGMLKNKTRILVITHGSIYLPQVDVII 840

QY 840 VMSGKISEMGSYQELLARDGAFAPFLRTYASTEQDABENGVTGVSGPGKEAKQWENG 899  
 DB 841 VMSGKISEMGSYQELLARDGAFAPFLRTYASTEQDABENGVTGVSGPGKEAKQWENG 900

QY 900 MLVTDGAGLQRLQSSSSSYSGDISRHNNSTAELOKAEAKKEETWKLMEADKAQTGVYK 959  
 DB 901 ILVTDGAGLQRLQSSSSSYSGDISRHNNSTAELOKAEAKKEETWKLMEADKAQTGVYK 959

QY 960 LSVTDWYKAIGLFISFLSIFLFCMNVHVSALASNYWLSLWTD - PIVNGTQEHKTVRLSV 1018  
 DB 960 LSVYNNYKAIGLCISFLSIFLFCMNVHVSALASNYWLSLWTD - PIVNGTQEHKTVRLSV 1019

QY 1019 YGALGISQIAVFGYSMAVSIIGGILASRLHVDLLHSLRSPMSFFERTPSGNLVNRFSK 1078  
 DB 1020 YGALGISQIAVFGYSMAVSIIGGILASRLHVDLLHSLRSPMSFFERTPSGNLVNRFSK 1079

QY 1079 ELDTVDSMIPBIVKMFMSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSR 1138  
 DB 1080 ELDTVDSMIPBIVKMFMSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSR 1139

QY 1139 QLKRLSVSRSPVSHFNETLLGVSVIRAFBEQERFIHQSDLKVDENQKAYYPSIVANRW 1198  
 DB 1140 QLKRLSVSRSPVSHFNETLLGVSVIRAFBEQERFIHQSDLKVDENQKAYYPSIVANRW 1199

QY 1199 LAVRLCVCNCIVLFAALFAVIRSHLSAGLVGSVSLQVTTVTLNWLVRMSSMETNI 1258  
 DB 1200 LAVRLCVCNCIVLFAALFAVIRSHLSAGLVGSVSLQVTTVTLNWLVRMSSMETNI 1259

QY 1259 VAVRLKEYSETEKEAPWQIQTAPPSPWPQVGRVEFRNYCLRYREDLDFVLRHINVTIN 1318  
 DB 1260 VAVRLKEYSETEKEAPWQIQTAPPSPWPQVGRVEFRNYCLRYREDLDFVLRHINVTIN 1319

QY 1319 GGEKVGIVGRTGAKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPV 1378  
 DB 1320 GGEKVGIVGRTGAKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPV 1379

QY 1379 LFSGLRNLDLPPSOYSDSEEVWTSLELAHLKDFVSALPKLDHCEACGENLSVQROLV 1438  
 DB 1380 LFSGLRNLDLPPSOYSDSEEVWTSLELAHLKDFVSALPKLDHCEACGENLSVQROLV 1439

QY 1439 CLARALLRKTILVLDEATAAVDLETDLLIOSTIRTOPEDCVTLTIAHRLNTIMDYTRVI 1498  
 DB 1440 CLARALLRKTILVLDEATAAVDLETDLLIOSTIRTOPEDCVTLTIAHRLNTIMDYTRVI 1499

QY 1499 VLDKGEIQEYCAPSLLQOQGLFYFSMAKDAGLV 1531  
 DB 1500 VLDKGEIQEYCAPSLLQOQGLFYFSMAKDAGLV 1532

RESULT 12  
 Q8CG09 PRELIMINARY; PRT; 1532 AA.  
 AC Q8CG09;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Multidrug resistance-associated protein 1.  
 GN Name=Mrp1;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.



RC	STRAIN=Sprague-Dawley; TISSUE=Brain;	
RA	Yang Z., Li C.S.W., Shen D.L., Ho R.J.Y.;	
RL	Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.	
CC	1- SIMILARITY: Belongs to the ABC transporter family.	
DR	EMBL; AL170916; AN08532.1; -	
DR	HSSP; P08716; 1MTO.	
DR	GO; GO:0016021; C: integral to membrane; IEA.	
DR	GO; GO:0005524; F: ATP binding; IEA.	
DR	GO; GO:0043262; F: ATPase activity, coupled to transmembrane m. . . ; IEA.	
DR	GO; GO:0001666; F: nucleotide binding; IEA.	
DR	GO; GO:0005215; F: transporter activity; IEA.	
DR	GO; GO:0008610; P: transport; IEA.	
DR	InterPro; IPR003593; AAA ATPase.	
DR	InterPro; IPR011527; ABC membrane 1.	
DR	InterPro; IPR011440; ABC TM transp.	
DR	InterPro; IPR003439; ABC transporter.	
DR	InterPro; IPR001395; ABC/ket_red.	
DR	InterPro; IPR005292; MRP assoc.	
DR	Pfam; PF00664; ABC membrane; 2.	
DR	Pfam; PF00005; ABC tran; 2.	
DR	ProDom; PD000006; ABC transporter; 2.	
DR	SMART; SM00382; AAA; 2.	
DR	TIGRFAMs; TIGR00957; MRP assoc_pro; 1.	
DR	PROSITE; PS00929; ABC TMIF; 2.	
DR	PROSITE; PS00211; ABC TRANSPORTER_1; 2.	
DR	PROSITE; PS00893; ABC TRANSPORTER_2; 2.	
DR	PROSITE; PS00063; ALDO-KETO_REDUCTASE_3; UNKNOWN_1.	
KW	ATP-binding.	
SQ	SEQUENCE 1532 AA; 171557 MW; C18F3554DD85732A CRC64;	
	Query Match 70.0%; Score 6932.5; DB 2; Length 1532;	
	Best Local Similarity 87.3%; Pred. No. 0;	
	Matches 1338; Conservative 100; Mismatches 92; Indels 3; Gaps 3;	
QY	1 MALRFCGADGSDPLDNDNNVNTWNTSDPTKCFQNTLVWVPCFYLWACFPFFLYLSRH 60	
DB	1 MALRFCSSDSDPLDNDNNVNTWNTSDPTKCFQNTLVWVPCFYLWACFPFFLYLSRH 60	
QY	61 DRGYQMTPLNKTALGFLMWICWADLFYSFWERSRGIPLAPVFLVSPFLGTTLLA 120	
DB	61 DRGYQMTPLNKTALGFLMWICWADLFYSFWERSRGIPLAPVFLVSPFLGTTLLA 120	
QY	121 TFLQLRRKGVSIGIMTLFWLVALVCALAILRSKIMTALKEDAQVDFLFDITFYVFS 180	
DB	121 TFLQLRRKGVSIGIMTLFWLVALVCALAILRSKIMTALKEDAQVDFLFDITFYVFS 180	
QY	181 LLLQLVLSCFSDSPLETHDPNCPSSASFLSITITWITGLIVRGYRPLEGSD 240	
DB	181 LVLQLVLSCFSDSPLETHDPNCPSSASFLSITITWITGLIVRGYRPLEGSD 240	
QY	241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVWVS - SKDPAQPKSSKVDANEVEA 299	
DB	241 LWSLNKEDTSEVVPVLVKNWKECAKTRKQPVKVWVS - SKDPAQPKSSKVDANEVEA 299	
QY	300 LIVKSPQKEMNPSLFKVLKTFGPFYFMSFFKAIHDLMPGPOIKLLIKFVNDTKAP 359	
DB	301 LIVKSKHRRDPSLFKVLKTFGPFYFMSFFKAIHDLMPGPOIKLLIKFVNDTKAP 359	
QY	360 DWQGYFVTLVFTACIQLVHLYVHICFVSGWRIKTAIVGAVYRKALVITNSARKSST 419	
DB	361 DWQGYFVTLVFTACIQLVHLYVHICFVSGWRIKTAIVGAVYRKALVITNSARKSST 420	
QY	420 VGEIVNLSVDAQRFMDLATYINNIWSAPLQVILALYLLMLNLGSPVLAVGVMLVMPV 479	
DB	421 VGEIVNLSVDAQRFMDLATYINNIWSAPLQVILALYLLMLNLGSPVLAVGVMLVMPV 480	
QY	480 NAVNMTKTKTQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFKQKVLAIROBELKVL 539	
DB	481 NAVNMTKTKTQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFKQKVLAIROBELKVL 540	
QY	540 KKSAYLSAVGFTTWCTPFLVALCTFAVYVTTIDNNILDAQTAFLVSLALFNILRFPNLL 599	
DB	541 KKSAYLAAGFTTWCTPFLVALCTFAVYVTTIDNNILDAQTAFLVSLALFNILRFPNLL 600	
QY	600 PMWISSIVQASVSLKRLRIFLSHEELRPSDSTERRPVPKDGCGTNSITVRNATFTWARSDDP 659	
DB	601 PMWISSIVQASVSLKRLRIFLSHEELRPSDSTERRPVPKDGCGTNSITVRNATFTWARSDDP 660	
QY	660 TLNGITPSIPGALVAVVGVQVCGKSSLLSALLAEMDKVEGHVAKGSAVAVVPOQAWLQN 719	
DB	661 TLNGITPAIPGALVAVVGVQVCGKSSLLSALLAEMDKVEGHVAKGSAVAVVPOQAWLQN 720	
QY	720 DSLRENILFGCOLPEPYRSVIOACALLPDLIELIPSGDRTEIGKGNVLSGGQKQVSLA 779	
DB	721 DSLRENILFGCOLPEPYRSVIOACALLPDLIELIPSGDRTEIGKGNVLSGGQKQVSLA 780	
QY	780 RAVYSNADIVLFDPLSADVAHVKHIFENVIKPGMLKNKTRILVTHSMYSLPOVDVII 839	
DB	781 RAVYNSDIYLLDDPLSAVDAHVKHIFENVIKPGMLKNKTRILVTHSMYSLPOVDVII 840	
QY	840 VMGSKIEMSGSYOELLARDGAFARLTASTVEQEOADAEENGVTGSGPGKEAKOMENG 899	
DB	841 VMGSKIEMSGSYOELLARDGAFARLTASTVEQEOADAEENGVTGSGPGKEAKOMENG 900	
QY	900 MLVTSAGKQLQROLSSSSSYSGDISRHNSSTASLQKAEAKKETWKLMEADKAQTGVK 959	
DB	901 ILVTDVAVGKPLQRHLSNSSSHSVVNTQHGSTABLQKS - GVKEETWKLMEADKAQTGVK 959	
QY	960 LSVYWDYMKATGLFISLIFLFCNHNHVSALASNYWLSLWTD - PIVNGTQEHTKVRLSV 1018	
DB	960 LSVYWDYMKATGLFISLIFLFCNHNHVSALASNYWLSLWTD - PIVNGTQEHTKVRLSV 1019	
QY	1019 YGALGISQIAVFGYSMAVSGIGILASRCLFVLDLHLSILRSPMSFFERTSGNLVNRFSK 1078	
DB	1020 YGALGILQGVAVFGYSMAVSGIGIFASRLHLDLQNLVLRSPMSFFERTSGNLVNRFSK 1079	
QY	1079 ELDTVDSMIPEVIMKMGSLFNVIGACIVILLATPIAAIIIPPLGLYFFVQRYVASSR 1138	
DB	1080 ELDTVDSMIPEVIMKMGSLFNVIGAVIIILLATPIAAVIPPPLGLYFFVQRYVASSR 1139	
QY	1139 QLKRESVSRSPVYSHENETLLGVSVIRAFEEORRTHQSDLKVDENOKAYYPSIVANRW 1198	
DB	1140 QLKRESVSRSPVYSHENETLLGVSVIRAFEEORRTHQSDLKVDENOKAYYPSIVANRW 1199	
QY	1199 LAVRLECVGNCIVFAALFAVIRSHLSAGLVLSYSLSQVTTYLNLVLRMSSEMETNI 1258	
DB	1200 LAVRLECVGNCIVFAALFAVIRSHLSAGLVLSYSLSQVTTYLNLVLRMSSEMETNI 1259	
QY	1259 VAVERLKEYSETEKAPWQIQTAPSSWFOVGRVEFRNYCLRYREDLDFVLRHINTIN 1318	
DB	1260 VAVERLKEYSETEKASWQIQTAPSPWPHSGRVEFRDYCLRYREDLDFVLRHINTIE 1319	
QY	1319 GGEKVGIVGRTGAGKSLTLGLFRINESARGEIIIDGINIAKIGLHDLRKFITIIPODPV 1378	
DB	1320 GGEKVGIVGRTGAGKSLTLGLFRINESARGEIIIDGINIAKIGLHDLRKFITIIPODPV 1379	
QY	1379 LFGSLRMNLDPPFSQVSDSEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRQV 1438	
DB	1380 LFGSLRMNLDPPFSQVSDSEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRQV 1439	
QY	1439 CLARALLRKTKILVLDATVADLETTDLTQSTIRTOFEDCTVLTAAHRLNTIMDYTRVI 1498	
DB	1440 CLARALLRKTKILVLDATVADLETTDLTQSTIRTOFEDCTVLTAAHRLNTIMDYTRVI 1499	
QY	1499 VLDKGEIQEYCAPSDLIQORGLFVSNMAKDAGLV 1531	
DB	1500 VLDKGEIQEYCAPSDLIQORGLFVSNMAKDAGLV 1532	
	RESULT 13	
	Q810G9	
ID	Q810G9 PRELIMINARY; PRT; 1523 AA.	
AC	Q810G9; (Tremblrel. 24, Created)	
DT	01-JUN-2003 (Tremblrel. 24, Last sequence update)	
DT	01-JUN-2003 (Tremblrel. 26, Last annotation update)	



DE ATP-binding cassette protein C1 variant A.  
GN Name=Abcc1;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;  
RA Takayanagi S., Iehikawa T.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Belongs to the ABC transporter family.  
DR EMBL; AY174892; AA04983.1; -;  
DR HSSP; P08716; IMTO.  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0005524; F: ATP binding; IEA.  
DR GO; GO:0042626; F: ATPase activity, coupled to transmembrane m. . .; IEA.  
DR GO; GO:0000166; F: nucleotide binding; IEA.  
DR GO; GO:0005215; F: transporter activity; IEA.  
DR GO; GO:0006810; P: transport; IEA.  
DR InterPro; IPR003593; AAA ATPase.  
DR InterPro; IPR011527; ABC membrane 1.  
DR InterPro; IPR001140; ABC\_TM\_transp.  
DR InterPro; IPR003439; ABC\_transporter.  
DR InterPro; IPR001395; ABC\_transporter.  
DR InterPro; IPR005292; MRP\_assoc.  
DR Pfam; PF00664; ABC membrane; 2.  
DR Pfam; PF00005; ABC tran; 2.  
DR ProDom; PD000006; ABC transporter; 2.  
DR SMART; SM00382; AAA; 2.  
DR TIGRFAMs; TIGR00957; MRP\_assoc\_pro; 1.  
DR PROSITE; PS0929; ABC TM1F; 2.  
DR PROSITE; PS0211; ABC\_TRANSPORTER\_1; 2.  
DR PROSITE; PS0893; ABC\_TRANSPORTER\_2; 2.  
DR PROSITE; PS0063; ALDO\_KETO\_REDUCTASE\_3; UNKNOWN\_1.  
KW ATP-binding.  
SQ SEQUENCE 1523 AA; 170503 MW; B40337051A1CB9C6 CRC64;

Query Match 69.6%; Score 6892; DB 2; Length 1523;  
Best Local Similarity 87.0%; Pred. No. 0;  
Matches 1333; Conservative 98; Mismatches 90; Indels 12; Gaps 4;

QY 1 MALRGCSDGSDPLMDNNTWNTNPDFTKCFQNTLVWPCFYLWACFPFYLYLSRH 60  
DB 1 MALRGCSDGSDPLMDNNTWNTNPDFTKCFQNTLVWPCFYLWACFPFYLYLSRH 60

QY 61 DRGYIQMTPLNKTALGFLLWVWVADLFLYFWERSRGIPLAPFLVSPTLIGITLLA 120  
DB 61 DRGYIQMTPLNKTALGFLLWVWVADLFLYFWERSRGIPLAPFLVSPTLIGITLLA 120

QY 121 TFLIQLERRKGVQSSGIMLTFWLVALCALALRSKIMTALKEDAQVDLFRDITFYVYFS 180  
DB 121 TFLIQLERRKGVQSSGIMLTFWLVALCALALRSKIMTALKEDAQVDLFRDITFYVYFS 180

QY 181 LLLIQLVLCFSDRSLFSETIHDNPPCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
DB 181 LLLIQLVLCFSDRSLFSETIHDNPPCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240

QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYVS-SKDPAPKSSSKVDANEEVEA 299  
DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYVS-SKDPAPKSSSKVDANEEVEA 299

QY 300 LIVKSPQKWNPSLFVLYKTFGPFYFLMSFFPKALHDLMMFSGPQILLIKFVNDTKAP 359  
DB 300 LIVKSPQKWNPSLFVLYKTFGPFYFLMSFFPKALHDLMMFSGPQILLIKFVNDTKAP 359

QY 360 DWQGYFTVLLFVTLACQTLVHQYFHCIFVSGMRKTAIVGAVYRKALVITNSARKSST 419  
DB 360 DWQGYFTVLLFVTLACQTLVHQYFHCIFVSGMRKTAIVGAVYRKALVITNSARKSST 419

QY 420 VGEIVNLSVDAQRMDLATYINMTWSAPLQVTLALYLWNLGSPVLGAVVMTLMPVS 479  
DB 420 VGEIVNLSVDAQRMDLATYINMTWSAPLQVTLALYLWNLGSPVLGAVVMTLMPVS 479

QY 421 VGEIVNLSVDAQRMDLATYINMTWSAPLQVTLALYLWNLGSPVLGAVVMTLMPVS 480  
DB 421 VGEIVNLSVDAQRMDLATYINMTWSAPLQVTLALYLWNLGSPVLGAVVMTLMPVS 480

QY 480 NAVMAKTKTYQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQBELKVL 539  
DB 480 NAVMAKTKTYQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQBELKVL 539

QY 540 KKSAYLSAVGFTWVCTPFLVALCTFAVYVVTIDENNILDQAOTAFVSLALFNILRPLNLT 599  
DB 540 KKSAYLSAVGFTWVCTPFLVALCTFAVYVVTIDENNILDQAOTAFVSLALFNILRPLNLT 599

QY 600 PMWISSIVQASVSLKRLRIFLSHBELEPDSIERRPVKDGGGTNSITVRNATFTWARSDDP 659  
DB 600 PMWISSIVQASVSLKRLRIFLSHBELEPDSIERRSIKDGGGWNSTVKNATFTWARDDEP 660

QY 660 TLNGITFSIPGALVAVVGVQCGKSSLLSALLAEMDKVEGHVATKGSVAVVPPQAWION 719  
DB 660 TLNGITFSIPGALVAVVGVQCGKSSLLSALLAEMDKVEGHVATKGSVAVVPPQAWION 720

QY 720 DSLRENILFGQLEPEPYRSVIOACALLPDLEILPSGDRTEIGEGKVNLSGQKORVSLA 779  
DB 720 DSLRENILFGQLEPEPYRSVIOACALLPDLEILPSGDRTEIGEGKVNLSGQKORVSLA 780

QY 780 RAVYSNADIYLFDDPLSAVDHVGHIIPENVIGPKGMLKNKTRILVTHSMSYLPQVDVII 839  
DB 780 RAVYSNADIYLFDDPLSAVDHVGHIIPENVIGPKGMLKNKTRILVTHSMSYLPQVDVII 840

QY 840 VMSGKISEMGSYQELLDGCAFAEFRTYASTQEOBAENGVTGVSQPKGKAKOMENG 899  
DB 840 VMSGKISEMGSYQELLDGCAFAEFRTYASTQEOBAENGVTGVSQPKGKAKOMENG 900

QY 900 MLVTDGAKQLORQLSSSSSYSGDISRHNSHTAELOKAEKEETWKLMEADKAOTGVK 959  
DB 900 MLVTDGAKQLORQLSSSSSYSGDISRHNSHTAELOKAEKEETWKLMEADKAOTGVK 959

QY 960 LSVYDYMKAIGLFTISFLSIFLFCMCHVSALASNYWLSLWTDG-PVANGTQSHTKVRLSV 1018  
DB 960 LSVYDYMKAIGLFTISFLSIFLFCMCHVSALASNYWLSLWTDG-PVANGTQSHTKVRLSV 1019

QY 1019 YGALGISQGIADVGVSMVSTGGILASRCLHVDLLHSILRSPMSFFERTPGNLRNRSK 1078  
DB 1019 YGALGISQGIADVGVSMVSTGGILASRCLHVDLLHSILRSPMSFFERTPGNLRNRSK 1079

QY 1079 ELDTVDSMIPVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVYASSR 1138  
DB 1079 ELDTVDSMIPVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVYASSR 1139

QY 1139 QKRLSVSRPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENKAYYPSIVANRW 1198  
DB 1139 QKRLSVSRPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENKAYYPSIVANRW 1199

QY 1199 LAVRLECVGNCIVLFAALFAVIRSHSLSAGLVGLSVSYSLQITAYLNLVLRMSSEMTNI 1258  
DB 1199 LAVRLECVGNCIVLFAALFAVIRSHSLSAGLVGLSVSYSLQITAYLNLVLRMSSEMTNI 1259

QY 1259 VAVERLKEYSETEKAPWQIOETAPPSWPOQVRVEFRNYCLRYREDLDFVLRHINVTIN 1318  
DB 1259 VAVERLKEYSETEKAPWQIOETAPPSWPOQVRVEFRNYCLRYREDLDFVLRHINVTIN 1319

QY 1319 GGEKGVIGRTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTIIPQDPRV 1378  
DB 1319 GGEKGVIGRTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTIIPQDPRV 1379

QY 1379 LFSGLRNLDLPPFSQYSDDEEVTMSLELAHKDFVSALPDKLDHCEAEGENLSVGQRQLV 1438  
DB 1379 LFSGLRNLDLPPFSQYSDDEEVTMSLELAHKDFVSALPDKLDHCEAEGENLSVGQRQLV 1439

QY 1439 CLARALLKTKILVLDEATAVLDLTDLLIOSTIRTFQEDCTVLIHRLNTIMDYTRVI 1498  
DB 1439 CLARALLKTKILVLDEATAVLDLTDLLIOSTIRTFQEDCTVLIHRLNTIMDYTRVI 1499

QY 1499 VLDKGEIEYCAPSDLLQORGLFYSMKADAGLV 1531  
DB 1499 VLDKGEIEYCAPSDLLQORGLFYSMKADAGLV 1532

```

RESULT 14
Q68CP7 PRELIMINARY; PRT; 1215 AA.
ID Q68CP7;
AC Q68CP7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp781G125.
GN Name=DKFZp781G125;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amysdala;
RG The German cDNA Consortium;
RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Oeanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR749835; CAH10691.1; -.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001140; ABC TM transp.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR005292; MRP assoc.
DR InterPro; IPR000719; Prot kinase.
DR Pfam; PF00564; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Hypothetical protein.
SQ SEQUENCE 1215 AA; 134926 MW; E490AD1F5268F72E CRC64;

Query Match 61.3%; Score 6075; DB 2; Length 1215;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1202; Conservative 0; Mismatches 3; Indels 10; Gaps 1;

QY 327 MSFFPKAIDHLMFSGPQILKLIKFNVDNTRKAPDWQGYFYTVLLFVTAQLOTLVLHGVFF 386
DB 1 MSFFPKAIDHLMFSGPQILKLIKFNVDNTRKAPDWQGYFYTVLLFVTAQLOTLVLHGVFF 60

QY 387 ICFVSGMEIKTAVIGAVYRKALVITNSARKSTVGEIYNLMSVDAQRFMDLATYINMWS 446
DB 61 ICFVSGMEIKTAVIGAVYRKALVITNSARKSTVGEIYNLMSVDAQRFMDLATYINMWS 120

QY 447 APLQVILALYLLNLGPSVLGAVVYLVMPVNAVMAKTKTYQVAHMKSKDNRIKLMN 506
DB 121 APLQVILALYLLNLGPSVLGAVVYLVMPVNAVMAKTKTYQVAHMKSKDNRIKLMN 180

QY 507 BILNGIKVLKYAMELAFKDLAIROBELKVLKKSAYLSAVGTFTWCTPFLVALCTFA 566
DB 181 BILNGIKVLKYAMELAFKDLAIROBELKVLKKSAYLSAVGTFTWCTPFLVALCTFA 240

QY 567 VVVTIDENNIIDAQTAFLSLALENLRLPILPWLVISSIVQASVSLKRLIFLSHEELE 626
DB 241 VVVTIDENNIIDAQTAFLSLALENLRLPILPWLVISSIVQASVSLKRLIFLSHEELE 300

QY 627 PDSIERRPVKDGGTNSITVRNATFTWARSDDPTLNGITFTSIPGALVAVVGQVCGKSS 686
DB 301 PDSIERRPVKDGGTNSITVRNATFTWARSDDPTLNGITFTSIPGALVAVVGQVCGKSS 360

QY 687 LLSALLAEMDKVEGHVAIKGSVAVVPOQAWTQNDLSRENILFGQLEPEYRVSIVQACAL 746
DB 361 LLSALLAEMDKVEGHVAIKGSVAVVPOQAWTQNDLSRENILFGQLEPEYRVSIVQACAL 420

QY 747 LPDLIELPSGDRTEIGEGVNLGGOKORVSLARAVYNSADIYLFDDPLSAVDAHVGKHI 806
DB 421 LPDLIELPSGDRTEIGEGVNLGGOKORVSLARAVYNSADIYLFDDPLSAVDAHVGKHI 480

```

```

QY 807 FENVIGPKMLKNKTRILVTHSMYSYLPQVDVIIVMGGKISMGSGYQELLARDGAFBFL 866
DB 481 FENVIGPKMLKNKTRILVTHSMYSYLPQVDVIIVMGGKISMGSGYQELLARDGAFBFL 540

QY 867 RTYASTEEOQDAEEN-----GVTVSGPGKEAKQMGMLVTDGAKQLOQLSS 916
DB 541 RTYASTEEOQDAEENGVTVSGPGKEAKQMGMLVTDGAKQLOQLSS 600

QY 917 SSSYSGDISRHHNSTAEQKAEKKEETKWLMEADKAQTQGVKLSVYWDYMKALGLFTSF 976
DB 601 SSSYSGDISRHHNSTAEQKAEKKEETKWLMEADKAQTQGVKLSVYWDYMKALGLFTSF 660

QY 977 LSIPLFMCHVSNVSLASNYWLSLWTDPIVNGTQDHTKVLRLSVYGALGISQGIAPVGYGMA 1036
DB 661 LSIPLFMCHVSNVSLASNYWLSLWTDPIVNGTQDHTKVLRLSVYGALGISQGIAPVGYGMA 720

QY 1037 VSIGGILASRCLHVDLLHLSILRSPMSFPTPSGNLVNRFPSKELDTVDSDMIPVTKMPMG 1096
DB 721 VSIGGILASRCLHVDLLHLSILRSPMSFPTPSGNLVNRFPSKELDTVDSDMIPVTKMPMG 780

QY 1097 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQKRLLESVSRSPVYSHFN 1156
DB 781 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQKRLLESVSRSPVYSHFN 840

QY 1157 ETLIGVSVIRAFERQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNICVILFAAL 1216
DB 841 ETLIGVSVIRAFERQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNICVILFAAL 900

QY 1217 FAVISRHSLSAGLVGLSVSLQVTVYLVNLMVRMSMETNIVAVERLKEYSETEKAPW 1276
DB 901 FAVISRHSLSAGLVGLSVSLQVTVYLVNLMVRMSMETNIVAVERLKEYSETEKAPW 960

QY 1277 QIQTAPPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGGKGVIGVGTGAGKSSL 1336
DB 961 QIQTAPPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGGKGVIGVGTGAGKSSL 1020

QY 1337 TLGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLFSGSLRNWLDPFQSYSD 1396
DB 1021 TLGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLFSGSLRNWLDPFQSYSD 1080

QY 1397 EEVWTSLELAHLKDFVSALPKDHECAEGENLSVGQRLVCLARALLRKTILVLDEA 1456
DB 1081 EEVWTSLELAHLKDFVSALPKDHECAEGENLSVGQRLVCLARALLRKTILVLDEA 1140

QY 1457 TAAVDLETDDLIOSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1516
DB 1141 TAAVDLETDDLIOSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1200

QY 1517 QRGIFYSMAXDAGLV 1531
DB 1201 QRGIFYSMAXDAGLV 1215

RESULT 15
Q80ZK8 PRELIMINARY; PRT; 1519 AA.
ID Q80ZK8;
AC Q80ZK8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to ATP-binding cassette, sub-family C (CFTR/MRP), member 3 (Fragment).
DE (Fragment).
GN Name=Abcc3;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

```

-!- SIMILARITY: Belongs to the ABC transporter family.

CC EMBL; BCO48825; AAH48825.1; -.  
 DR HSP; P26361; 1R0Z.  
 DR MGD; MG1:1923658; Abcc3.  
 DR GO; GO:0005887; C-integral to plasma membrane; IDA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR011527; ABC membrane 1.  
 DR InterPro; IPR001140; ABC TM transpt.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR005292; MRP assoc.  
 DR Pfam; PF00664; ABC membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD00006; ABC transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAMs; TIGR00957; MRP assoc pro; 1.  
 DR PROSITE; PS00929; ABC TMIF; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 KW ATP-binding.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 1519 AA; 168648 MW; 72A5151A862EC948 CRC64;  
 Query Match 45.7%; Score 4522; DB 2; Length 1519;  
 Best Local Similarity 56.2%; Pred. No. 2.5e-270;  
 Matches 865; Conservative 269; Mismatches 366; Indels 38; Gaps 8;  
 QY 8 SADGSDPLMDNVTWNTGNPDTKCFQNTVLVWVPCFLWACFPFFLYLGRHGRGYIQM 67  
 DB 4 SGELGSKFWDNLSIYTTPTDTPCFQNSLLAWPCIVLWALPCYFLRHHQGIYVL 63  
 QY 68 TPLNKTATGALLMIVCWADLFYSFWERSRGIFLAPVPLVPTLLGTTLLATLFIQLE 127  
 DB 64 SWLSRLKLTALGVLLWCVSWDLFYFHLGHGSSPAPVFFVPLVVGITMLLATLLIQYE 123  
 QY 128 RRGVQSSGIMLTFMLVALCALILRSKIMTALKEDAQVLDLFRDITVYVFSLLLIQIV 187  
 DB 124 RLRGVQSSGVLIIFLLCQIAIIPFRSKILSALAEKILDPFRFTTYIIPALVFCALI 183  
 QY 188 LSCFSDRPLFSETHDNPCESSASFLSRITFWITGLIVRGYRQLEGSDLWSLNKE 247  
 DB 184 LSCFKEPLPSPENLDINPCPEASAGPFRSLFWFTRFALIGYRRPLEDRDLWSLSE 243  
 QY 248 DTSEQVPLVKNWKECAKTRKQPVVYVSKDPAQPKESKVDANEVEBALIVKSPQK 307  
 DB 244 DCSHKVQRLLEAMQKQ---QNASGSGTATAEPKIPGE-----DAVLLKPRPK 289  
 QY 308 EWNPSLFVLYTFCGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPDWQGYFT 367  
 DB 290 SKQSPSLRALVTRFTSSLLMSACFNLIQNLGFPNPQLLSILIRFISDPTAPTWGFLLA 349  
 QY 368 VLLFYTACLQTLVHQYFHCIFVSGMRKIKTAVIGAYRKALVITNSARKSSTVGEIVNLM 427  
 DB 350 GLMFSSITMQLILHQYHCFVWALRLRTALIGVYRKALVITNSVKRESTVGEVNL 409  
 QY 428 SVDAQRFMDLATYINMINSAPQVILALYLWNLGSPVLAGVAVMLVMPVNAVMAKMT 487  
 DB 410 SVDAQRFMDVSPFINLLSAPQVILALYLWNLGSPVLAGVAVVILLPLNGAVSMKM 469  
 QY 488 KTYQVAHMKSKDNRIKLWNEILNGIKVLAYAWELAPKDKVLAIQBELKVLKSAIYLSA 547  
 DB 470 KTYQVKQKFKDSRIKLWSEILNGIKVLKLYAWEPSPLEQVKGIRQSELQRLKAGYLQA 529  
 QY 548 VGTFTWCTPPLVALCTFAVYVITDENLIDQATFVSLAFNLIRPPLNIPMVISIV 607  
 DB 530 ISTFTWICTPPLVLTITLVGVYVDESINVLDEKAFVSLFNLILKIPLNMLPQISGLT 589  
 QY 608 QASVSLKRLRIFLSHEEPEPSIERRPVKDGGGTSITVRNATFTWARSDDPTLNGITFS 667  
 DB 590 QASVSLKRIQDFLNQELDPQCVKRTISPG---YAIITHNGTFTWAQDLRPTLHSLNIQ 646  
 QY 668 IPEGALVAVGVGCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQWQONDSIRENIL 727  
 DB 647 IPKGAIVAVGVGCGKSLVALLGEMEKLEGVVSVKGSVAVYVQQAQWQONCTIQENVL 706

QY 728 FGQLEEFYRSVIAQACALLPDLEILPSGDRTEIGKGNVLSGGQKRVSLARAVYSNAD 787  
 DB 707 FGQPMNPKRYQOALETCAADLDVLPGGDQTEIGKGINLSGGQKRVSLARAVYSNAD 766  
 QY 788 IYLDDEPLSADVAHVGHKIFENVIGPKGMLKNKTRILYTHSMXYLPQVDVIVMGGKIS 847  
 DB 767 IFLDDPLSADVSHVAKHIFDQVIGPEGVAGKTRVLVTHGISFLPQDFDIIVLAGGQVS 826  
 QY 848 EMGSYQELLARDGAPAFELRYASTEQDQAEENGVTGVSFGPKAEKQWENGML-----VT 903  
 DB 827 EMGHYSALLQHDGSAFANPLRYAPDEQDEHE-----ALQANAEVLLLEDLTSLTHDIT 881  
 QY 904 DS-----AGKQLOQLSSSSSYSGDIS-----RHNSTABEQKAEKEETWKLMEADK 952  
 DB 882 DNEPAIVEVRKQFMREMSLS--EGEVQNRTPMKKHTNSLEKEALVTTKETGALKEBI 940  
 QY 953 AOTQVKLSVVDYWKAGLIFSLFPLFCNHNVSALASNYWLSLWTDVINGTQHT 1012  
 DB 941 AETGNVLSVVDYAKSMGLCTTILSICLLYGGQSAANAIGANVWLSANSNDABEHQOQNK 1000  
 QY 1013 KVLRSVYGALGISQGIAGVFGYSMAVISGILASRCLHVDLHLSILRSPMSFFERTPSGNL 1072  
 DB 1001 SVRLGVTAALGILQGLLWLSAFTWVGAIQARLLHEALLHNKIRSPQSFDTTPSGRI 1060  
 QY 1073 VNRFSKELDTYDSMIPEVIKMFMSLNFVIGACIVILLATPIAAIIIPPLGLIYFVQRF 1132  
 DB 1061 LNRFSKDIYIDEVLAPTILMLNSFFTSISTIMVIVASTPLEVMVVLPLAVLYGFVQRF 1120  
 QY 1133 YVASRQLKLESVRSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDKLDVENQKAYPS 1192  
 DB 1121 YVATSRQLKLESISRSPISFHSFSETVTGTSVIRAYGRIOQDFKVLSDTKVDNNQKSSYP 1180  
 QY 1193 IVANRWLAVRLECVGNCIVLFAALFAVLSRHSLSAGLVGLSVYSLOVTTVYLNWLVRMSS 1252  
 DB 1181 IASNRWLGHVHVFVGNVCVLFALFAVIGRNSLAPGLVGLSVYALQVTMALNWMIRMS 1240  
 QY 1253 EMETNIVAVERLKEYSETEKEAPWQIQETAPPSPWPQVGRVYFRNYCLRYREDLDFVLRH 1312  
 DB 1241 DLESNIIVAVERVKEYSKTKTEAPVWVESNRAPEGWPTGRGWYFRNYSVRYRPGLEVLKN 1300  
 QY 1313 INVINGEKVIGVGRGAGKSSITGLFRINESAGEIIIDGINIAKIGLHDLRPKITI 1372  
 DB 1301 VTVHVGGEKVGIVGRGAGKSSMTCLFRILEAAGEIVIDGLMVAHIGLHDLRSQITI 1360  
 QY 1373 IPQDPVLPFSGSLRWNLDPFSDYSDVEVWTSLELAHLKDFVSALPKLDHECAEGGENLSV 1432  
 DB 1361 IPQDPILFSGTLRWNLDPFGYSEDIWRALSHLNTFVSQPAGLDFQCAEGGDNLSV 1420  
 QY 1433 GQQLVCLARALLRKTKILVLDATAAVIDLETDLLIQSTIRTFQFEDCTVLTIAHRLNTIM 1492  
 DB 1421 GQQLVCLARALLRKSRVLDEATAIDLETDLLIQSTIRTFQFEDCTVLTIAHRLNTIM 1480  
 QY 1493 DYTRVILDKGEIYEGYAPSDLLQORGLFYSMADAGL 1530  
 DB 1481 DYNRVILDKGVAAEFDSFVNLIAAGGIFYGMKADAGL 1518

Search completed: March 18, 2005, 11:12:16  
 Job time : 212.942 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:20 ; Search time 40.7091 Seconds  
(without alignments)  
3570.253 Million cell updates/sec

Title: US-10-665-283-8

Perfect score: 10016

Sequence: 1 MALRGFCSADGSDPLMDWNV.....RSVAVAKRPFISPDLS 1947

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/1/aaa/5A\_COMB.pep:\*
- 2: /cgn2\_6/prodata/1/aaa/5B\_COMB.pep:\*
- 3: /cgn2\_6/prodata/1/aaa/6A\_COMB.pep:\*
- 4: /cgn2\_6/prodata/1/aaa/6B\_COMB.pep:\*
- 5: /cgn2\_6/prodata/1/aaa/PCFUS\_COMB.pep:\*
- 6: /cgn2\_6/prodata/1/aaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	78.5	1531	1	US-08-463-092B-4
2	7860	78.5	1531	2	US-08-462-109A-4
3	7860	78.5	1531	3	US-08-460-907B-4
4	7860	78.5	1531	3	US-08-463-179A-4
5	7860	78.5	1531	3	US-08-461-384B-4
6	7860	78.5	1531	4	US-09-647-140B-19
7	7849	78.4	1531	1	US-08-141-893-2
8	7849	78.4	1531	1	US-08-463-092B-2
9	7849	78.4	1531	2	US-08-462-109A-2
10	7849	78.4	1531	3	US-08-460-907B-2
11	7849	78.4	1531	3	US-08-463-179A-2
12	7849	78.4	1531	3	US-08-461-384B-2
13	7849	78.4	1531	3	US-08-407-207A-2
14	7002.5	69.9	1528	1	US-08-463-092B-6
15	7002.5	69.9	1528	2	US-08-462-109A-6
16	7002.5	69.9	1528	3	US-08-460-907B-6
17	7002.5	69.9	1528	3	US-08-463-179A-6
18	7002.5	69.9	1528	3	US-08-461-384B-6
19	4487.5	44.8	1527	4	US-09-647-140B-6
20	4487.5	44.8	1530	4	US-09-647-140B-33
21	3391.5	33.9	1503	4	US-09-647-140B-8
22	3385.5	33.8	1503	4	US-09-792-616-3
23	3318.5	32.1	1498	4	US-09-792-616-9
24	2374.5	23.7	1621	3	US-08-972-927-3
25	2349.5	23.5	1622	3	US-08-972-927-6
26	2266.5	22.6	1325	4	US-09-647-140B-2
27	2195.5	21.9	1261	3	US-09-439-313-538

28	2195.5	21.9	1261	4	US-09-636-215-538	Sequence 538, App
29	2195.5	21.9	1261	4	US-09-685-166A-538	Sequence 538, App
30	2195.5	21.9	1261	4	US-09-679-426-538	Sequence 538, App
31	2195.5	21.9	1261	4	US-09-759-143-538	Sequence 538, App
32	2195.5	21.9	1261	4	US-09-651-236-538	Sequence 538, App
33	2118	21.1	1581	3	US-08-726-320-3	Sequence 3, Appli
34	2118	21.1	1581	3	US-09-208-716-3	Sequence 3, Appli
35	2115	21.1	1228	3	US-09-439-313-537	Sequence 537, App
36	2115	21.1	1228	4	US-09-636-215-537	Sequence 537, App
37	2115	21.1	1228	4	US-09-685-166A-537	Sequence 537, App
38	2115	21.1	1228	4	US-09-679-426-537	Sequence 537, App
39	2115	21.1	1228	4	US-09-759-143-537	Sequence 537, App
40	2115	21.1	1228	4	US-09-651-236-537	Sequence 537, App
41	2082	20.8	1437	3	US-09-061-400-2	Sequence 2, Appli
42	2082	20.8	1453	3	US-09-001-273-2	Sequence 2, Appli
43	2082	20.8	1453	3	US-08-843-459A-2	Sequence 38, Appli
44	2081	20.8	1437	4	US-10-162-012-38	Sequence 4, Appli
45	2080	20.8	1437	4	US-09-647-140B-4	

## ALIGNMENTS

RESULT 1  
US-08-463-092B-4  
; Sequence 4, Application US/08463092B  
; Patent No. 5766880  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
; STREET: Queen's University at Kingston  
; CITY: Kingston  
; STATE: Ontario  
; COUNTRY: CANADA  
; ZIP: K7L 3N6  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,092B  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steeg, Carol Miernicki  
; REGISTRATION NUMBER: 39,539  
; REFERENCE/DOCKET NUMBER: Q1546  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (613) 545-2342  
; TELEFAX: (613) 545-6853  
; INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:	
LENGTH: 1531 amino acids	
TYPE: amino acid	
TOPOLOGY: linear	
MOLECULE TYPE: protein	
US-08-463-092B-4	
Query Match	78.5%; Score 7860; DB 1; Length 1531;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1531; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MALRGFCSADGSDPLMDWNVNTWNTSNPDKFQNTLVVWPCFYLMACFPFYFLYLSRH 60
DB	1 MALRGFCSADGSDPLMDWNVNTWNTSNPDKFQNTLVVWPCFYLMACFPFYFLYLSRH 60
QY	61 DRGYQMTPLNKTATGALLMTVCWADLFYFWERSRGIFLAPVFLVSPDLLGITLLA 120
DB	61 DRGYQMTPLNKTATGALLMTVCWADLFYFWERSRGIFLAPVFLVSPDLLGITLLA 120
QY	121 TELIQLERRKGVOSSGIMLTFWLVALCALAIRSKIMTALKEDAQVDLFRDITFYVYFS 180
DB	121 TELIQLERRKGVOSSGIMLTFWLVALCALAIRSKIMTALKEDAQVDLFRDITFYVYFS 180
QY	181 LLLIQLVSCFSDRSPLETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
DB	181 LLLIQLVSCFSDRSPLETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
QY	241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKOPVYVSKDPAQPKESKVDANEVEAL 300
DB	241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKOPVYVSKDPAQPKESKVDANEVEAL 300
QY	301 IVKSPQKEWNPFLKVLKTYFGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
DB	301 IVKSPQKEWNPFLKVLKTYFGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
QY	361 WQGYFYTVLLFVTAQLTLVHLQYHICFVSGMRKTAIVGAVYKALVITNSARKSSTV 420
DB	361 WQGYFYTVLLFVTAQLTLVHLQYHICFVSGMRKTAIVGAVYKALVITNSARKSSTV 420
QY	421 GEIVNLSVDAORFMDLATYINMWSAPLOVTLALYLWLNLPSPVLAGVAVMLVMPVN 480
DB	421 GEIVNLSVDAORFMDLATYINMWSAPLOVTLALYLWLNLPSPVLAGVAVMLVMPVN 480
QY	481 AVAMKTKTYQVAHMKSDNRKIKLMEILNGIKVLKYAWELAFKDKVLAIROELKVLK 540
DB	481 AVAMKTKTYQVAHMKSDNRKIKLMEILNGIKVLKYAWELAFKDKVLAIROELKVLK 540
QY	541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDNNILDAQTAFAVSLALFNILRPLNLP 600
DB	541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTDNNILDAQTAFAVSLALFNILRPLNLP 600
QY	601 MVISSIVQASVSLKRLRIFLSEELPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660
DB	601 MVISSIVQASVSLKRLRIFLSEELPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660
QY	661 LNTFTSIPEGALVAVVQVCGKSSLSALLAEMDKVEGHVAIKGSVAYVPQAWIQND 720
DB	661 LNTFTSIPEGALVAVVQVCGKSSLSALLAEMDKVEGHVAIKGSVAYVPQAWIQND 720
QY	721 SLRENILFGCOLLEPYRSVTOACALLPDLILPSGDRTEIGEKNVLSGGQKORVSLAR 780
DB	721 SLRENILFGCOLLEPYRSVTOACALLPDLILPSGDRTEIGEKNVLSGGQKORVSLAR 780
QY	781 AVYSNADILYFDDPLSADVAHVGHIFENVTGPKGMLKNKTRILVTHSMSVLPQVDVILV 840
DB	781 AVYSNADILYFDDPLSADVAHVGHIFENVTGPKGMLKNKTRILVTHSMSVLPQVDVILV 840
QY	841 MSGGKISEMGVQELLARDAFAFLRYASTEQEQAENGVTGVSQPGKEAKOMENGM 900
DB	841 MSGGKISEMGVQELLARDAFAFLRYASTEQEQAENGVTGVSQPGKEAKOMENGM 900
QY	901 LVTSAGKQLORQLSSSSSSSGDISRRHNSAEQLQKAEKKEETWKLMEADKATQGVKL 960
DB	901 LVTSAGKQLORQLSSSSSSSGDISRRHNSAEQLQKAEKKEETWKLMEADKATQGVKL 960
QY	961 SVYDYMKAIGLFIISFLSIFLPMCHVNSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020
DB	961 SVYDYMKAIGLFIISFLSIFLPMCHVNSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020
QY	1021 ALGISQGIATVFGYSMAVSGIGGILASRCLVHVDLLHSILSPMSFFPERTSGNLVNRFSKEL 1080
DB	1021 ALGISQGIATVFGYSMAVSGIGGILASRCLVHVDLLHSILSPMSFFPERTSGNLVNRFSKEL 1080
QY	1081 DTVDMSIPEVIMKPMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140
DB	1081 DTVDMSIPEVIMKPMGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140
QY	1141 KRLSVSRSPYSHFNETHLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200
DB	1141 KRLSVSRSPYSHFNETHLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200
QY	1201 VRLECVGNICIVLFAALPAVISRHSLSAGLVGLSVSYSLQVTTYLNWLVRMSSEMETNIVA 1260
DB	1201 VRLECVGNICIVLFAALPAVISRHSLSAGLVGLSVSYSLQVTTYLNWLVRMSSEMETNIVA 1260
QY	1261 VERLKEYSETEKEAPWQIQETAPPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTNGG 1320
DB	1261 VERLKEYSETEKEAPWQIQETAPPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTNGG 1320
QY	1321 EKVGIIVGTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRPFKTIIPQDPVLF 1380
DB	1321 EKVGIIVGTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRPFKTIIPQDPVLF 1380
QY	1381 SGSLRMNLDPFSQVSDSEVWTSLELAHKDFVSALPKLDHECAEGGENLSVGQRLVCL 1440
DB	1381 SGSLRMNLDPFSQVSDSEVWTSLELAHKDFVSALPKLDHECAEGGENLSVGQRLVCL 1440
QY	1441 ARALLRTKILVLDEATAAVDLETDLLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIL 1500
DB	1441 ARALLRTKILVLDEATAAVDLETDLLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIL 1500
QY	1501 DKGSIQEVGAPSDLLQORGLFYSMAXDAGLV 1531
DB	1501 DKGSIQEVGAPSDLLQORGLFYSMAXDAGLV 1531
RESULT 2	
US-08-462-109A-4	
; Sequence 4, Application US/08462109A	
; Patent No. 5882875	
; GENERAL INFORMATION:	
; APPLICANT: Cole, Susan P.C.	
; APPLICANT: Deeley, Roger G.	
; TITLE OF INVENTION: METHODS FOR IDENTIFYING	
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS	
; NUMBER OF SEQUENCES: 6	
; CORRESPONDENCE ADDRESS:	
; ADDRESSEE: LAHIVE & COCKFIELD	
; STREET: 60 State Street, suite 510	
; CITY: Boston	
; STATE: Massachusetts	
; COUNTRY: USA	
; ZIP: 02109	
; COMPUTER READABLE FORM:	
; MEDIUM TYPE: Floppy disk	
; COMPUTER: IBM PC compatible	
; OPERATING SYSTEM: PC-DOS/MS-DOS	
; SOFTWARE: ASCII text	
; CURRENT APPLICATION DATA:	
; APPLICATION NUMBER: US/08/462,109A	
; FILING DATE:	
; CLASSIFICATION: 435	
; PRIOR APPLICATION DATA:	
; APPLICATION NUMBER: 07/966,923	
; FILING DATE: 27-OCT-1992	
; APPLICATION NUMBER: 08/029,340	

```

; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: PQI-002CP4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-462-109A-4

Query Match 78.5%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSDGSDPLMDNMTNTSNPDKCFQNTLVWPCFYLWACFPFFYLYLSRH 60
DB 1 MALRGFCSDGSDPLMDNMTNTSNPDKCFQNTLVWPCFYLWACFPFFYLYLSRH 60
QY 61 DRGYTQMTPLNKTALGFLMWVWADLFYFWSRSGIFLAPVFLVSPVTLGTTLLA 120
DB 61 DRGYTQMTPLNKTALGFLMWVWADLFYFWSRSGIFLAPVFLVSPVTLGTTLLA 120
QY 121 TFLQLERKGVQSSGIMTLFWLVALCALALIRSKIMTALKEDAQVDLFRDITFVYVFS 180
DB 121 TFLQLERKGVQSSGIMTLFWLVALCALALIRSKIMTALKEDAQVDLFRDITFVYVFS 180
QY 181 LLLIQLVLSFCSDRSPLESETIHDNPPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVLSFCSDRSPLESETIHDNPPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKYVSSKDPAPKESKVDANEVEAL 300
DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKYVSSKDPAPKESKVDANEVEAL 300
QY 301 IVKSPQKWNPSLFVLYKTFPGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
DB 301 IVKSPQKWNPSLFVLYKTFPGPYFLMSPFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
QY 361 WQGYFYTVLLFVYACLTQLTLVHLHQYFHCIFVSGMRKTAIVAGVYRKALVITNSARKSTV 420
DB 361 WQGYFYTVLLFVYACLTQLTLVHLHQYFHCIFVSGMRKTAIVAGVYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRFDMDLATYINMIWSAPLOVILALYLLMLNIGPSVLGAVVWLMVPVN 480
DB 421 GEIVNLSVDAQRFDMDLATYINMIWSAPLOVILALYLLMLNIGPSVLGAVVWLMVPVN 480
QY 481 AVAMKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540
DB 481 AVAMKTKTYQVAHMKSKDNRIKLWNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540
QY 541 KSAYLSAVGTFTVCTPFLVALCTFAVYVITDENNILDQAQAFVSLALFNILRFFLNILP 600
DB 541 KSAYLSAVGTFTVCTPFLVALCTFAVYVITDENNILDQAQAFVSLALFNILRFFLNILP 600
QY 601 MYISSIVQASVSLKRLRIFLSHEELEPPDSIERPPKVGCGGTSITVRNATFTWASDPPT 660
DB 601 MYISSIVQASVSLKRLRIFLSHEELEPPDSIERPPKVGCGGTSITVRNATFTWASDPPT 660
QY 661 LNGITFISIEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIOND 720
DB 661 LNGITFISIEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIOND 720
QY 721 SLRENILFCQLEEPYRSVIAQCALLPDLEILPSGDRTEIGEKGYNLSGGQKQVSLAR 780

```

```

DB 721 SLRENILFCQLEEPYRSVIAQCALLPDLEILPSGDRTEIGEKGYNLSGGQKQVSLAR 780
QY 781 AVYSNADIYLPDDPLSADVAHVGHKHPENVIGPKGMLKNKTRILVTHSMYLPQVDVIV 840
DB 781 AVYSNADIYLPDDPLSADVAHVGHKHPENVIGPKGMLKNKTRILVTHSMYLPQVDVIV 840
QY 841 MSGGKISEMGYSQELLARDGAFAEFLRTYASTEQQDAEENGVTGSGPGKEAKWENGM 900
DB 841 MSGGKISEMGYSQELLARDGAFAEFLRTYASTEQQDAEENGVTGSGPGKEAKWENGM 900
QY 901 LVTDSAGKOLORQLSSSSSYSGDTSRHNSHTAELOKAKAEKKEETWKLHEADQAQGVKL 960
DB 901 LVTDSAGKOLORQLSSSSSYSGDTSRHNSHTAELOKAKAEKKEETWKLHEADQAQGVKL 960
QY 961 SVYDYMKAIGLFTISFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTVLSVYG 1020
DB 961 SVYDYMKAIGLFTISFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTVLSVYG 1020
QY 1021 ALGISQGIAGVGYSMASVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSEL 1080
DB 1021 ALGISQGIAGVGYSMASVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSEL 1080
QY 1081 DTVDSMIDPEVIMFMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140
DB 1081 DTVDSMIDPEVIMFMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140
QY 1141 KRLESVSRSPPVSHFNETHLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200
DB 1141 KRLESVSRSPPVSHFNETHLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200
QY 1201 VRLECVGNCIVLFAALFAVIRSHSLSAGLVAGLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260
DB 1201 VRLECVGNCIVLFAALFAVIRSHSLSAGLVAGLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETKEAPWQIQETAPPSSWPQGRVFEFRNYCLRYREDLDFVLRHINTYINGG 1320
DB 1261 VERLKEYSETKEAPWQIQETAPPSSWPQGRVFEFRNYCLRYREDLDFVLRHINTYINGG 1320
QY 1321 EKVGVGTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGVGTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380
QY 1381 SGSLRMNLDPPFSQYSDSEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440
DB 1381 SGSLRMNLDPPFSQYSDSEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440
QY 1441 ARALLRKTILVLEATAAVIDETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
DB 1441 ARALLRKTILVLEATAAVIDETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIQEYGAPESSDILLQORGLFYSMADAGLV 1531
DB 1501 DKGEIQEYGAPESSDILLQORGLFYSMADAGLV 1531

RESULT 3
US-08-460-907B-4
; Sequence 4, Application US/08460907B
; Patent No. 5891724
; GENERAL INFORMATION:
; APPLICANT: Deelev, Roger G.
; APPLICANT: Cole, Susan P.C.
; TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
; RESISTANCE ON A CELL
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6

```





Db 1501 DKGEIQEYCAPSDDLQOQGLFYSMADAGLV 1531

RESULT 4  
US-08-463-179A-4  
; Sequence 4, Application US/08463179A  
; Patent No. 6001563  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,179A  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeConti, Giulio A. Jr.  
; REGISTRATION NUMBER: 31,503  
; REFERENCE/DOCKET NUMBER: PQ1-002CP8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1531 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-463-179A-4

Query Match 78.5%; Score 7860; DB 3; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MALRGCSDGSDPLMDNMVNTWNTSNPDTKCFQNTVLVWPCFYLMACFPFYLYLSRH 60  
Db 1 MALRGCSDGSDPLMDNMVNTWNTSNPDTKCFQNTVLVWPCFYLMACFPFYLYLSRH 60  
QY 61 DRGYIQMTPLNKTALGFLLMVVCWADLFYFWSRSGIFLAPVFLVSPPTLLGITTLLA 120  
Db 61 DRGYIQMTPLNKTALGFLLMVVCWADLFYFWSRSGIFLAPVFLVSPPTLLGITTLLA 120  
QY 121 TELIQLERRKGVQSSGIMLTFWLVALVCALALILRSKIMTALKEDAQVDLFRDITFYVPS 180  
Db 121 TELIQLERRKGVQSSGIMLTFWLVALVCALALILRSKIMTALKEDAQVDLFRDITFYVPS 180  
QY 181 LLLIQLVLCFSDRSPFLSETIHDNPNCPSSASFLSRITTFWITGLIVRGYRQPLEGSD 240  
Db 181 LLLIQLVLCFSDRSPFLSETIHDNPNCPSSASFLSRITTFWITGLIVRGYRQPLEGSD 240  
QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKRESSKVDANBEVEAL 300

Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKRESSKVDANBEVEAL 300  
QY 301 IVKSPQKEWNPDLFKVLYKTPGYPFLMSFFPKAIIHDLMMFSGPQILKLIKFVNDTKAPD 360  
Db 301 IVKSPQKEWNPDLFKVLYKTPGYPFLMSFFPKAIIHDLMMFSGPQILKLIKFVNDTKAPD 360  
QY 361 WQGYFTYVLLFVTAQLOTLVLHQYFPHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
Db 361 WQGYFTYVLLFVTAQLOTLVLHQYFPHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
QY 421 GEIVNLMGSDAQRFPMDLATYINMIWSAPLOVLALYLLWNLGSPVLAGVAVMLVMPVN 480  
Db 421 GEIVNLMGSDAQRFPMDLATYINMIWSAPLOVLALYLLWNLGSPVLAGVAVMLVMPVN 480  
QY 481 AVMAKMTKTYOVAHMKSNDRIKLMNEILNGIKVLKYLAWELAFKDKVLAIQBELKVLK 540  
Db 481 AVMAKMTKTYOVAHMKSNDRIKLMNEILNGIKVLKYLAWELAFKDKVLAIQBELKVLK 540  
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
QY 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
Db 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
QY 661 LMGITTFIPEGALVAVVGQVCGKSSLSALLAEMDKVEGHVAIKGSVAYVPPQQAOWND 720  
Db 661 LMGITTFIPEGALVAVVGQVCGKSSLSALLAEMDKVEGHVAIKGSVAYVPPQQAOWND 720  
QY 721 SLRENILFGCQLEEPYRSVIAQACALLPDLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
Db 721 SLRENILFGCQLEEPYRSVIAQACALLPDLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
QY 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKMKLNKTRILVTHSMYSYLPQVDVIV 840  
Db 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKMKLNKTRILVTHSMYSYLPQVDVIV 840  
QY 841 MSGGKISEMGVYQELLARDGAPAEFLRTYASTEQEDAEENGVTGSGPGKAKOMENGM 900  
Db 841 MSGGKISEMGVYQELLARDGAPAEFLRTYASTEQEDAEENGVTGSGPGKAKOMENGM 900  
QY 901 LVTDGAGKQLOQLSSSSSSSYSGDISRHHNSTAELQKAEAKKEETWKLMEADKAQTQGVKL 960  
Db 901 LVTDGAGKQLOQLSSSSSSSYSGDISRHHNSTAELQKAEAKKEETWKLMEADKAQTQGVKL 960  
QY 961 SVYDYMKAIGLFTISFLIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
Db 961 SVYDYMKAIGLFTISFLIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
QY 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHLSILRSPMSFFERTPSGNLVNRFSEKL 1080  
Db 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHLSILRSPMSFFERTPSGNLVNRFSEKL 1080  
QY 1081 DTVDSMIEVIMKFMGSLFNIVIGACIVILLATPAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
Db 1081 DTVDSMIEVIMKFMGSLFNIVIGACIVILLATPAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
QY 1141 KRLESVSRSVPYSHNETLLGVSVIRAPEEQRFIHQSDLKVDENQKAYPSIVANRWLA 1200  
Db 1141 KRLESVSRSVPYSHNETLLGVSVIRAPEEQRFIHQSDLKVDENQKAYPSIVANRWLA 1200  
QY 1201 VRLECVGNCIVLFAALFAVISRHSISAGLVGSYSIQVTTYLNWLVNRMSEMETNIVA 1260  
Db 1201 VRLECVGNCIVLFAALFAVISRHSISAGLVGSYSIQVTTYLNWLVNRMSEMETNIVA 1260  
QY 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVEPRNYCLRYREDLDFVLRHINTYINGG 1320  
Db 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVEPRNYCLRYREDLDFVLRHINTYINGG 1320  
QY 1321 EKGVIGVGTGACKSGLTLGLFRINSEAGEIIIDGINIAKIGLHDLRPFKTIIPQDVLV 1380

Db 1321 EKVGIVGRTGAGKSSLTGLPRINESAGEIIIDGINIAKIGLHDLRFXKTIIPQDPVLF 1380  
Qy 1381 SGLRNLDLPFSQYSDDEEVTWTSLELAHLKDFVSALPKLDHCECAGGENLSVGQQLVCL 1440  
Db 1381 SGLRNLDLPFSQYSDDEEVTWTSLELAHLKDFVSALPKLDHCECAGGENLSVGQQLVCL 1440  
Qy 1441 ARALLRKTKILVLDATAVDETDLDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Db 1441 ARALLRKTKILVLDATAVDETDLDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Qy 1501 DKGEIQEYGAPSDLLQORGLFYSMADAGLV 1531  
Db 1501 DKGEIQEYGAPSDLLQORGLFYSMADAGLV 1531

## RESULT 5

US-08-461-384B-4  
; Sequence 4, Application US/08461384B  
; Patent No. 6025473  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
; STREET: Queen's University at Kingston  
; CITY: Kingston  
; STATE: Ontario  
; COUNTRY: CANADA  
; ZIP: K7L 3N6  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,384B  
; FILING DATE: 05-JUN-95  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steeg, Carol Miernicki  
; REGISTRATION NUMBER: 39,539  
; REFERENCE/DOCKET NUMBER: Q1547  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (613) 545-2342  
; TELEFAX: (613) 545-6853  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1531 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-461-384B-4

Query Match 78.5%; Score 7860; DB 3; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MALRGFCSDGSDPLWDMNVTWNTSNPDFTKCFQNTVLVWVPCFVLWACFPFFFLYLSRH 60  
Db 1 MALRGFCSDGSDPLWDMNVTWNTSNPDFTKCFQNTVLVWVPCFVLWACFPFFFLYLSRH 60  
Qy 61 DRGIQMTPLNKTALGFLWIVCWADLFYSPWERSRGIFLAPVFLVSPILLGITLLA 120

Db 61 DRGIQMTPLNKTALGFLWIVCWADLFYSPWERSRGIFLAPVFLVSPILLGITLLA 120  
Qy 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDITFYVYFVS 180  
Db 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDITFYVYFVS 180  
Qy 181 LLLIQLVLSCSFSDSPFLSETIHDNPNCPSSASFLSRTTFWMTGLVRCVROPLEGSD 240  
Db 181 LLLIQLVLSCSFSDSPFLSETIHDNPNCPSSASFLSRTTFWMTGLVRCVROPLEGSD 240  
Qy 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANEVEAL 300  
Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANEVEAL 300  
Qy 301 IVKSPQKEWNPSPFKVLYKTFPGPYFLMSFFPKAIDHLMFSGPQILKLIIFVNDTKAPD 360  
Db 301 IVKSPQKEWNPSPFKVLYKTFPGPYFLMSFFPKAIDHLMFSGPQILKLIIFVNDTKAPD 360  
Qy 361 WQGYFYTVLLFVTACLOTLVLHOYFHICFVSGMRKTAIVGAVYRKALAVITNSARKSSTV 420  
Db 361 WQGYFYTVLLFVTACLOTLVLHOYFHICFVSGMRKTAIVGAVYRKALAVITNSARKSSTV 420  
Qy 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLMLNLPSPSVLAGVAVMLVPEVN 480  
Db 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLMLNLPSPSVLAGVAVMLVPEVN 480  
Qy 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYAWELAFKDKVLAIROBELKVLK 540  
Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYAWELAFKDKVLAIROBELKVLK 540  
Qy 541 KSAVLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFVLSLAFNLIRPLNLTP 600  
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFVLSLAFNLIRPLNLTP 600  
Qy 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWARSDDPT 660  
Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWARSDDPT 660  
Qy 661 LNGITFSIPGALVAVVGVCGCKSSLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIQND 720  
Db 661 LNGITFSIPGALVAVVGVCGCKSSLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIQND 720  
Qy 721 SLRENILFGCOLEBPYRSVIQACALLPDLLEILPSGDRTEIGEKNVLSGQKQKORVSLAR 780  
Db 721 SLRENILFGCOLEBPYRSVIQACALLPDLLEILPSGDRTEIGEKNVLSGQKQKORVSLAR 780  
Qy 781 AVYSNADIYLFDDPLSAVDAHVGKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVIV 840  
Db 781 AVYSNADIYLFDDPLSAVDAHVGKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVIV 840  
Qy 841 MSGGKISEMGSYQELLARDGNAFAEFRTYASTEQOEQDAENGVTGVSQPGKEAKOMENGM 900  
Db 841 MSGGKISEMGSYQELLARDGNAFAEFRTYASTEQOEQDAENGVTGVSQPGKEAKOMENGM 900  
Qy 901 LVTDSAGKQLQRLSSSSSSSGDLSRHNSHNSAEIQAEEKAEETWKLMEADKAQTQGVKL 960  
Db 901 LVTDSAGKQLQRLSSSSSSSGDLSRHNSHNSAEIQAEEKAEETWKLMEADKAQTQGVKL 960  
Qy 961 SVYWDYMKAI GLFTISFLSIFL FMCNHNVSALASNWLSTLWTDPIVNGTQEHKTVRLSVYG 1020  
Db 961 SVYWDYMKAI GLFTISFLSIFL FMCNHNVSALASNWLSTLWTDPIVNGTQEHKTVRLSVYG 1020  
Qy 1021 ALGISQGIAGVFGYSMAVSIIGIILASRCLHVDLHLSILRSPMSFPERTPSGNLVNRFSEKL 1080  
Db 1021 ALGISQGIAGVFGYSMAVSIIGIILASRCLHVDLHLSILRSPMSFPERTPSGNLVNRFSEKL 1080  
Qy 1081 DTVDSDMPEVIMKPMGSLFNVICACIVILLATPIAAIIIPPLGLIYFFVQFFYVASSRQL 1140  
Db 1081 DTVDSDMPEVIMKPMGSLFNVICACIVILLATPIAAIIIPPLGLIYFFVQFFYVASSRQL 1140  
Qy 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
Db 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200

QY 1201 VRLECVGNCIVLFAALFAVIRSHSLSAGLVSVSYSLQVTTYNLWLRMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVIRSHSLSAGLVSVSYSLQVTTYNLWLRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKEAPWQIETAPPSSWPQVGRVFRNCLRYREDLDFVLRHINTVINGG 1320  
 DB 1261 VERLKEYSETEKEAPWQIETAPPSSWPQVGRVFRNCLRYREDLDFVLRHINTVINGG 1320  
 QY 1321 EKVGIIVGRTGACKSSITLGLFRINSAGEEIIIDGINIAKIGHDLRFKTIIPODPVLF 1380  
 DB 1321 EKVGIIVGRTGACKSSITLGLFRINSAGEEIIIDGINIAKIGHDLRFKTIIPODPVLF 1380  
 QY 1381 SGLSRMLNDPFPQYDEEVTLSLAHLKDFVSALPDKLDHECARGENLSVGQRLVCL 1440  
 DB 1381 SGLSRMLNDPFPQYDEEVTLSLAHLKDFVSALPDKLDHECARGENLSVGQRLVCL 1440  
 QY 1441 ARALLRKTKILVLDATAVLETDLQSTTRTQFEDCTVTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ARALLRKTKILVLDATAVLETDLQSTTRTQFEDCTVTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEYGAPSDLLQORGLFYSMADAGLV 1531  
 DB 1501 DKGEIQEYGAPSDLLQORGLFYSMADAGLV 1531

RESULT 6  
 US-09-647-140B-19  
 ; Sequence 19, Application US/09647140B  
 ; Patent No. 6803184  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fox Chase Cancer Center  
 ; APPLICANT: Kruh, Gary D.  
 ; APPLICANT: Lee, Kun  
 ; APPLICANT: Belinsky, Martin G.  
 ; APPLICANT: Bain, Lisa J.  
 ; TITLE OF INVENTION: MRP-Related ABC Transporter Encoding  
 ; FILE REFERENCE: FCCC 98-02  
 ; CURRENT APPLICATION NUMBER: US/09/647,140B  
 ; PRIOR FILING DATE: 2001-05-21  
 ; PRIOR APPLICATION NUMBER: PCT/US99/06644  
 ; PRIOR FILING DATE: 1999-03-26  
 ; PRIOR APPLICATION NUMBER: 60/079,759  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/095,153  
 ; PRIOR FILING DATE: 1998-08-03  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 19  
 ; LENGTH: 1531  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapiens  
 US-09-647-140B-19

Query Match 78.5%; Score 7860; DB 4; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLWDMNVTWNTGNPDKCFQNTVLVWPCFYLWACFPFYLYLSRH 60  
 DB 1 MALRGFCSADGSDPLWDMNVTWNTGNPDKCFQNTVLVWPCFYLWACFPFYLYLSRH 60  
 QY 61 DRGYTQMTPLNKTALGELLIVCWADLFYFWSRSGIFLAPVPLVSPILLGTTLLA 120  
 DB 61 DRGYTQMTPLNKTALGELLIVCWADLFYFWSRSGIFLAPVPLVSPILLGTTLLA 120  
 QY 121 TFLIOLERRKGVQSSGIMLTFWLVLCALALIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TFLIOLERRKGVQSSGIMLTFWLVLCALALIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 QY 181 LLLIQLVLSCFSDRSPFSETHDPNCPPESSASFLSRITFWITGLIVRGVRLPQEGSD 240

DB 181 LLLIQLVLSCFSDRSPFSETHDPNCPPESSASFLSRITFWITGLIVRGVRLPQEGSD 240  
 QY 241 LWSLNKEDTSQVVPVLVKNWKKCEAKTRKQPVKVYVSSKQPAQPKSSKVDANEVEAL 300  
 DB 241 LWSLNKEDTSQVVPVLVKNWKKCEAKTRKQPVKVYVSSKQPAQPKSSKVDANEVEAL 300  
 QY 301 IVKSPQKEWNSLKFVLYKTFPGYFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 DB 301 IVKSPQKEWNSLKFVLYKTFPGYFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTTACLOTLVHOFPHICFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420  
 DB 361 WQGYFYTVLLFVTTACLOTLVHOFPHICFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLSMDVDAQRFMDLATYINMIWSAPLOVILALYLLNLMLGPSVLGAVVMVLMVFN 480  
 DB 421 GEIVNLSMDVDAQRFMDLATYINMIWSAPLOVILALYLLNLMLGPSVLGAVVMVLMVFN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIRQEBLKVVK 540  
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIRQEBLKVVK 540  
 QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFVLSLALFNILRFPNLI 600  
 DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFVLSLALFNILRFPNLI 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRRPVKGGTNSITVRNATFTWASDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRRPVKGGTNSITVRNATFTWASDPT 660  
 QY 661 LINGTFSIPEGALVAVVGQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQWIND 720  
 DB 661 LINGTFSIPEGALVAVVGQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQWIND 720  
 QY 721 SLRENILFGCQLEBEPYRSVIOACALLPDLLELPSGDRTEIGEKNVLSGGQKQVSLAR 780  
 DB 721 SLRENILFGCQLEBEPYRSVIOACALLPDLLELPSGDRTEIGEKNVLSGGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIKPGKMLKNKTRILVTHSMSYLPQVDVIV 840  
 DB 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIKPGKMLKNKTRILVTHSMSYLPQVDVIV 840  
 QY 841 MSGGKISMGSYQELLARDGAFELRTYASTEQDQAEENGVTGVSFGKEAKQWENGM 900  
 DB 841 MSGGKISMGSYQELLARDGAFELRTYASTEQDQAEENGVTGVSFGKEAKQWENGM 900  
 QY 901 LVTDSAGLQORQLSSSSSSSGDISRHHNSTAELQKAEKKEETWKLMEADKAQTGVKL 960  
 DB 901 LVTDSAGLQORQLSSSSSSSGDISRHHNSTAELQKAEKKEETWKLMEADKAQTGVKL 960  
 QY 961 SVYDYMKAIGLFTISFLIFMCMNHVSALASNYWLSLWTDPIVNGTQEHKVRLSVYG 1020  
 DB 961 SVYDYMKAIGLFTISFLIFMCMNHVSALASNYWLSLWTDPIVNGTQEHKVRLSVYG 1020  
 QY 1021 ALGISQGIAGVGYSMVSGIGLASCRLHVDLLHLSILSPMSFFERTSGNLVNRFSKEL 1080  
 DB 1021 ALGISQGIAGVGYSMVSGIGLASCRLHVDLLHLSILSPMSFFERTSGNLVNRFSKEL 1080  
 QY 1081 DTVDMSIPEVIMKMGSLFNIVIGACIVILATPIAAIIIPPLGLIYFFVQRFVASSROL 1140  
 DB 1081 DTVDMSIPEVIMKMGSLFNIVIGACIVILATPIAAIIIPPLGLIYFFVQRFVASSROL 1140  
 QY 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVIRSHSLSAGLVSVSYSLQVTTYNLWLRMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVIRSHSLSAGLVSVSYSLQVTTYNLWLRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKEAPWQIETAPPSSWPQVGRVFRNCLRYREDLDFVLRHINTVINGG 1320  
 DB 1261 VERLKEYSETEKEAPWQIETAPPSSWPQVGRVFRNCLRYREDLDFVLRHINTVINGG 1320

QY 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEHIIIDGINIAKIGLHDLRFKIIIPQDPVLF 1380  
 DB 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEHIIIDGINIAKIGLHDLRFKIIIPQDPVLF 1380  
 QY 1381 SGLRNLNLPFFSQYDEEVMTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL 1440  
 DB 1381 SGLRNLNLPFFSQYDEEVMTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL 1440  
 QY 1441 ARALLKTKILVLDEATAVDLETDLIQSTIRTOPEDCVTLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ARALLKTKILVLDEATAVDLETDLIQSTIRTOPEDCVTLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEYGAPSDLLQORGLFYSMKADGLV 1531  
 DB 1501 DKGEIQEYGAPSDLLQORGLFYSMKADGLV 1531

RESULT 7  
 US-08-141-893-2  
 ; Sequence 2, Application US/08141893  
 ; Patent No. 5489519  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cole, Susan P.C.  
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEIN  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD  
 ; STREET: 60 STATE STREET, SUITE 510  
 ; CITY: BOSTON  
 ; STATE: MASSACHUSETTS  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; OPERATING SYSTEM: IBM PC compatible  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/141,893  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/966,923; 08/029,340  
 ; FILING DATE: 27-OCT-1992; 8-MAR-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Deconti, Giulio A. Jr.  
 ; REGISTRATION NUMBER: 31,503  
 ; REFERENCE/DOCKET NUMBER: PQI-002  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5149  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1531 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-141-893-2

Query Match 78.4%; Score 7849; DB 1; Length 1531;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MALRGFCSDGSDPLNDWNNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60  
 DB 1 MALRGFCSDGSDPLNDWNNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60  
 QY 61 DRGIQMTPLNKTALGFLWVCWADLFYSFWRSGRGIFLAPVFLVSPFTLIGITLLA 120  
 DB 61 DRGIQMTPLNKTALGFLWVCWADLFYSFWRSGRGIFLAPVFLVSPFTLIGITLLA 120

QY 121 TFLIOLERRKGVSSGIMLTFWLVALVCAIALIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TFLIOLERRKGVSSGIMLTFWLVALVCAIALIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 QY 181 LLLIOLVLSGSDRSPFLSETHDNPCESSASFLSRITFWITGLTVRGVRQPLEGSD 240  
 DB 181 LLLIOLVLSGSDRSPFLSETHDNPCESSASFLSRITFWITGLTVRGVRQPLEGSD 240  
 QY 241 LWSLNKEDTSQGVVPLVKNWKECAKTRKQPVKVYSSKDPAPKESKVDANEVEAL 300  
 DB 241 LWSLNKEDTSQGVVPLVKNWKECAKTRKQPVKVYSSKDPAPKESKVDANEVEAL 300  
 QY 301 IVKSPQKEWNSLKFVLYKTFPGPYFLMSFFPKAHDLMFSGPQILKLLIKFVNDTKAPD 360  
 DB 301 IVKSPQKEWNSLKFVLYKTFPGPYFLMSFFPKAHDLMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFTYVLLFVTTACLOTVLHQLVHFICFVSGMRIKTAIVIGAVYRKALVITNSARKSSTV 420  
 DB 361 WQGYFTYVLLFVTTACLOTVLHQLVHFICFVSGMRIKTAIVIGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLNLNLPSPVLGAVMVLMPVN 480  
 DB 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLNLNLPSPVLGAVMVLMPVN 480  
 QY 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEITLNGIKVLYKLYAWELAFKDKVLAIROBELKVLK 540  
 DB 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEITLNGIKVLYKLYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVVVTIDENNILDQAOTAFVLSLAFNLIRPLNLIP 600  
 DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVVVTIDENNILDQAOTAFVLSLAFNLIRPLNLIP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEBLEPDSITERRPVKGGGTSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEBLEPDSITERRPVKGGGTSITVRNATFTWARSDDPT 660  
 QY 661 LNGITFSIPEGALVAVVGVGCGKSSLSALLAEMDKVEGHVALKGSVAVVPOQAWTND 720  
 DB 661 LNGITFSIPEGALVAVVGVGCGKSSLSALLAEMDKVEGHVALKGSVAVVPOQAWTND 720  
 QY 721 SLRENILFGQLEPEPYRYSVIQACALLPDLLEILPSGDRTEIGEKNVLSGQKQKORVSLAR 780  
 DB 721 SLRENILFGQLEPEPYRYSVIQACALLPDLLEILPSGDRTEIGEKNVLSGQKQKORVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDARHVGKHIENVIKPKGMLKNKTRILVTHSMSVLPQVDVILV 840  
 DB 781 AVYSNADIYLFDDPLSAVDARHVGKHIENVIKPKGMLKNKTRILVTHSMSVLPQVDVILV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFABFLRTYASTEODAEENGVTGVSQPGKEAKQEMNGM 900  
 DB 841 MSGGKISEMGSYQELLARDGAFABFLRTYASTEODAEENGVTGVSQPGKEAKQEMNGM 900  
 QY 901 LVTDGAKQLQRLSSSSSYSGDISRHNSSTAELOKAEKKEETWKLMEADKAQTQGVKL 960  
 DB 901 LVTDGAKQLQRLSSSSSYSGDISRHNSSTAELOKAEKKEETWKLMEADKAQTQGVKL 960  
 QY 961 SVYWDYMKALGLFISFLSIFLFCNHNVSALASNYLSLWTDTPVNGTQSHTKVRLSVYG 1020  
 DB 961 SVYWDYMKALGLFISFLSIFLFCNHNVSALASNYLSLWTDTPVNGTQSHTKVRLSVYG 1020  
 QY 1021 ALGISQGIAGVFGYSMAVSIIGLILASRCLHVDLHLSILRSPMSFFERTPSGNLVRFSKEL 1080  
 DB 1021 ALGISQGIAGVFGYSMAVSIIGLILASRCLHVDLHLSILRSPMSFFERTPSGNLVRFSKEL 1080  
 QY 1081 DTVDSMIPEVVKMFMSGLFNVIIGACIVILLATPIAAIIIPPLGLIYFFVQRFVYASSRQL 1140  
 DB 1081 DTVDSMIPEVVKMFMSGLFNVIIGACIVILLATPIAAIIIPPLGLIYFFVQRFVYASSRQL 1140  
 QY 1141 KRLESVRSRSPVYSHENETLLGVSVIRAFEOERFIHQSDLKVDENOKAVYPSIVANRWLA 1200  
 DB 1141 KRLESVRSRSPVYSHENETLLGVSVIRAFEOERFIHQSDLKVDENOKAVYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNMLVRMSSEMETNIVA 1260

```

1201 VRLECGNCVILPAALFAVIRSHLSAGLVGLSVSYSLQVTTYNLWLRMSMETNIVA 1260
1261 VERLKEYSETEKAPQIOETAPSSWPQVGRVFRNRYCLRVREDLDVFLRHINVTINGG 1320
1261 VERLKEYSETEKAPQIOETAPSSWPQVGRVFRNRYCLRVREDLDVFLRHINVTINGG 1320
1261 VERLKEYSETEKAPQIOETAPSSWPQVGRVFRNRYCLRVREDLDVFLRHINVTINGG 1320
1321 EKVGIVGRTGAGKSSITLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
1321 EKVGIVGRTGAGKSSITLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
1381 SGLRNLNLPFQSYDEEYVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL 1440
1381 SGLRNLNLPFQSYDEEYVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL 1440
1441 ARALLRKTILVLDATAVDTLTDLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVL 1500
1441 ARALLRKTILVLDATAVDTLTDLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVL 1500
1501 DKGEIOEYGAPSDLLQOQGLFYSMADAGLV 1531
1501 DKGEIOEYGAPSDLLQOQGLFYSMADAGLV 1531

```

RESULT 8

```

US-08-463-092B-2
; Sequence 2, Application US/08463092B
; Patent No. 5766880
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Dealey, Roger G.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,092B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1546
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853

```

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-092B-2

Query Match 78.4%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGCFSADGSDPLMDNNTNTNPNPTKCFQNTLVWPCFYLVWPCFYLVWPCFYLVSRH 60
DB 1 MALRGCFSADGSDPLMDNNTNTNPNPTKCFQNTLVWPCFYLVWPCFYLVWPCFYLVSRH 60

QY 61 DRGYIQMTPLNKTALGFLMLWCVADLIFYFWERSRGIFLAPVFLVSPDLLGITLLA 120
DB 61 DRGYIQMTPLNKTALGFLMLWCVADLIFYFWERSRGIFLAPVFLVSPDLLGITLLA 120

QY 121 TFLIQLEERRKGVQSSGIMLTFWLVALCALALIRSKIMTALKEDAQVDLFRDITFYVYFS 180
DB 121 TFLIQLEERRKGVQSSGIMLTFWLVALCALALIRSKIMTALKEDAQVDLFRDITFYVYFS 180

QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240

QY 241 LWSLNKEDTSEQVVPVLYKNWKECAKTRKQPVKVYSSKDPQAPKESKVDANEVEAL 300
DB 241 LWSLNKEDTSEQVVPVLYKNWKECAKTRKQPVKVYSSKDPQAPKESKVDANEVEAL 300

QY 301 IVKSPQKEWNPSPKLVLYKTPGYPFLMSFFPKAIHDLMMFSGPQILKLIKFNVDTKAPD 360
DB 301 IVKSPQKEWNPSPKLVLYKTPGYPFLMSFFPKAIHDLMMFSGPQILKLIKFNVDTKAPD 360

QY 361 WQGYFYTVLLFVTACLTQLVLHQYFHCIFVSGMRITKAVIGAVYRKALVITNSARKSSTV 420
DB 361 WQGYFYTVLLFVTACLTQLVLHQYFHCIFVSGMRITKAVIGAVYRKALVITNSARKSSTV 420

QY 421 GEIVNLSVDAQRFMDLATYINMWSAPLOVTLALYLLNLNIGPSVAGVAVMLVMPVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMWSAPLOVTLALYLLNLNIGPSVAGVAVMLVMPVN 480

QY 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIQEEELKVLK 540
DB 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIQEEELKVLK 540

QY 541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTDENNILDAQTAFVSLALFNILRFPNLILP 600
DB 541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTDENNILDAQTAFVSLALFNILRFPNLILP 600

QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDDPT 660

QY 661 LMGITFSPISPEGALVAVGVGCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720
DB 661 LMGITFSPISPEGALVAVGVGCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720

QY 721 SLRENILFGCQLEEPYYSVIOACALLPDLEILPSGDRTEIGEKGVLNSGGQKQVSLAR 780
DB 721 SLRENILFGCQLEEPYYSVIOACALLPDLEILPSGDRTEIGEKGVLNSGGQKQVSLAR 780

QY 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKMLKNKTRILVTHSHSYLPQVDVIV 840
DB 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKMLKNKTRILVTHSHSYLPQVDVIV 840

QY 841 MSGGKISEMGSYQELLARDGAFABFLRTYASTEQQDAEENGVTGSGPKAKOMENGM 900
DB 841 MSGGKISEMGSYQELLARDGAFABFLRTYASTEQQDAEENGVTGSGPKAKOMENGM 900

QY 901 LVTDSAGKQLQRLSSSSSSYSYSGDISRHHNSTAELQKAEAKKEETWKLMEADKAQTQGVKL 960

```

[illegible]

## RESULT 9

RESULTS 9  
US-08-462-109A-2

Sequence 2, Application US/08462109A

; Patent No. 5882875

; GENERAL INFORMATION:

APPLICANT: Cole, Susan P.C.

APPLICANT: Deeley, Roger G.

**TITLE OF INVENTION: METHODS FOR IDENTIFYING**

TITLE OF INVENTION: METHODS FOR IDENTIFYING MULTIDRUG RESISTANT TUMOR CELLS

; NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE &amp; COCKFIELD

STREET: 60 State Street, suite 510

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

```

;
; MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

```
; SOFTWARE: ASCII text
```

; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/462,109A

FILING DATE:

CLASSIFICATION: 435

;  
;  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/966,923

FILING DATE: 27-OCT-1992

; APPLICATION NUMBER: 08/029,340  
 ; FILING DATE: 8-MAR-1993  
 ; APPLICATION NUMBER: 08/141,893  
 ; FILING DATE: 26-OCT-1993  
 ; APPLICATION NUMBER: 08/407,207  
 ; FILING DATE: 20-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: DeConti, Giulio A. Jr.  
 ; REGISTRATION NUMBER: 31,503  
 ; REFERENCE/DOCKET NUMBER: PQI-002CP4  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5941  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1531 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-462-109A-2

Query Match	78.4%	Score 7849	DB 2	Length 1531
Best Local Similarity	99.9%	Pred. No. 0		
Matches 1529	Conservative	Mismatches 0	Indels 0	Gaps 0

Qy	1	MALRGFCSADGSDPLMDWNVTWNTSPDFTKCFQNTVVLVWPCFYLMWACPPFFFLYLSRH	60
Db	1	MALRGFCSADGSDPLMDWNVTWNTSPDFTKCFQNTVVLVWPCFYLMWACPPFFFLYLSRH	60
Qy	61	DRGYIOMTPLNKTKTALGFLFLWVCWADLFYSPWERSRGIFLAPVFLVSPFTLLGITTLLA	120
Db	61	DRGYIOMTPLNKTKTALGFLFLWVCWADLFYSPWERSRGIFLAPVFLVSPFTLLGITTLLA	120
Qy	121	TFLIQIERRKGVOSSGIMLTFWLVALVCAIAILRSKIMTALKEDAQVDLFRDITFYVYFS	180
Db	121	TFLIQIERRKGVOSSGIMLTFWLVALVCAIAILRSKIMTALKEDAQVDLFRDITFYVYFS	180
Qy	181	LLLIQLVLSCFSDRSPLFSETIHDNPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD	240
Db	181	LLLIQLVLSCFSDRSPLFSETIHDNPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD	240
Qy	241	LWSINKEDTSEQVVPVLVKNWKECAKTRQPKVYVYSSKPAQPKESSKVDANEVEAL	300
Db	241	LWSINKEDTSEQVVPVLVKNWKECAKTRQPKVYVYSSKPAQPKESSKVDANEVEAL	300
Qy	301	IVKSPOKENWPSLFKVLKYTFGYPFLMSFFFKAIHOLMMFSGQIILKLAIKFVNDTKAPD	360
Db	301	IVKSPOKENWPSLFKVLKYTFGYPFLMSFFFKAIHOLMMFSGQIILKLAIKFVNDTKAPD	360
Qy	361	WQGYFYTVLLFVTVACLOTLLVHLQYPHICFVSGMIRIKTAVIGAYYRKALVITNSARKSSTV	420
Db	361	WQGYFYTVLLFVTVACLOTLLVHLQYPHICFVSGMIRIKTAVIGAYYRKALVITNSARKSSTV	420
Qy	421	GEIVNLMSVDAQRFMDLATYINMINSAPQVITLALYLLNLNGLSPVLAVAVMVLMPVN	480
Db	421	GEIVNLMSVDAQRFMDLATYINMINSAPQVITLALYLLNLNGLSPVLAVAVMVLMPVN	480
Qy	481	AVMAMTKTYQVAHMKSKDNRIKLMEIILNGIKVLKYAWELAFKDKVLAIRQBELKVLK	540
Db	481	AVMAMTKTYQVAHMKSKDNRIKLMEIILNGIKVLKYAWELAFKDKVLAIRQBELKVLK	540
Qy	541	KSAYLSAVGTFTWVCTPFLVALCTFAVYVYTDENNILDAQTAFVSLALFNILRFPJNLILP	600
Db	541	KSAYLSAVGTFTWVCTPFLVALCTFAVYVYTDENNILDAQTAFVSLALFNILRFPJNLILP	600
Qy	601	MVSISSVQASVSLKRLRIFLSHEELPEPDSIERPPVKDGGGTNSITVRNATFTFWARSDPPT	660
Db	601	MVSISSVQASVSLKRLRIFLSHEELPEPDSIERPPVKDGGGTNSITVRNATFTFWARSDPPT	660
Qy	661	LNGITPISIEGALVAVGVQVCGCKSLLSALLAEMDKVEGHVAIKGSVAYVPOQAIQND	720
Db	661	LNGITPISIEGALVAVGVQVCGCKSLLSALLAEMDKVEGHVAIKGSVAYVPOQAIQND	720

```

QY 721 SIRENIFGCGLEEBEYRSVIOACALLPDLTILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780
Db 721 SIRENIFGCGLEEBEYRSVIOACALLPDLTILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780
QY 781 AVYSNADILYFDPLSADVAHVGKHFENVIGPKGMLKNTKRLIVTHSMSYLPQVDVLIIV 840
Db 781 AVYSNADILYFDPLSADVAHVGKHFENVIGPKGMLKNTKRLIVTHSMSYLPQVDVLIIV 840
QY 841 MSGGKISEMGYSQELLARDGAFELRTIVASTEQDQAEENGVTGVSQPGKEAKOMENGM 900
Db 841 MSGGKISEMGYSQELLARDGAFELRTIVASTEQDQAEENGVTGVSQPGKEAKOMENGM 900
QY 901 LVTDAGKQLQRLSSSSYSYSDISRHNSTAELOKAEAKKEETWKLMEADKAQGTQVKL 960
Db 901 LVTDAGKQLQRLSSSSYSYSDISRHNSTAELOKAEAKKEETWKLMEADKAQGTQVKL 960
QY 961 SVYDYMKAIGLFIPLSTFLPMCHNVHVSALASNYLSLWTDPIVNGTQEHKVLRSYVG 1020
Db 961 SVYDYMKAIGLFIPLSTFLPMCHNVHVSALASNYLSLWTDPIVNGTQEHKVLRSYVG 1020
QY 1021 ALGISQGIAGVGYSMAGVIGGILASRCLHVDLLHSILRSPMSFFERTPSGMLVNRFSKEL 1080
Db 1021 ALGISQGIAGVGYSMAGVIGGILASRCLHVDLLHSILRSPMSFFERTPSGMLVNRFSKEL 1080
QY 1081 DTVDMSIPEVIMKMGSLFNVTGACIVILLATPIAAIIIPGLLIYFFVQFYVASSRQL 1140
Db 1081 DTVDMSIPEVIMKMGSLFNVTGACIVILLATPIAAIIIPGLLIYFFVQFYVASSRQL 1140
QY 1141 KRESVSRSPVSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200
Db 1141 KRESVSRSPVSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200
QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVSVSYSLQVTTYLNWLVMSSEMETNIVA 1260
Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVSVSYSLQVTTYLNWLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPQIQTAPPSPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKAPQIQTAPPSPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
QY 1321 EKVGIIVGRTGAGKSLTGLPRINESAGEIIIDGINAKIGLHDLRKTIIIPQDPVLF 1380
Db 1321 EKVGIIVGRTGAGKSLTGLPRINESAGEIIIDGINAKIGLHDLRKTIIIPQDPVLF 1380
QY 1381 SGLSRWNLDPFQSYDDEEVTWTSLELAHLKDFVSALPKDLHDCEAGGENLSVGQRLVCL 1440
Db 1381 SGLSRWNLDPFQSYDDEEVTWTSLELAHLKDFVSALPKDLHDCEAGGENLSVGQRLVCL 1440
QY 1441 ARALLRKTILVLDATAAVDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 ARALLRKTILVLDATAAVDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIQEYCAPSDLLQORGLFYSMAKDAGLV 1531
Db 1501 DKGEIQEYCAPSDLLQORGLFYSMAKDAGLV 1531

```

```

RESULT 10
US-08-460-907B-2
; Sequence 2, Application US/08460907B
; Patent No. 5891724
; GENERAL INFORMATION:
; APPLICANT: Deeley, Roger G.
; APPLICANT: Cole, Susan P.C.
; TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
; TITLE OF INVENTION: RESISTANCE ON A CELL
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSES: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA

```

```

; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,907B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1551
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-460-907B-2

```

```

Query Match 78.4%; Score 7849; DB 2; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDWNVTWNTSNPDTTKCFQNTLVWVPCFYLMACFPFYLYLSRH 60
Db 1 MALRGFCSADGSDPLMDWNVTWNTSNPDTTKCFQNTLVWVPCFYLMACFPFYLYLSRH 60
QY 61 DRGYIQTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPTLGTTLLA 120
Db 61 DRGYIQTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPTLGTTLLA 120
QY 121 TFLQLERRKGVQSSGIMLTFWLVALVCALAILRSKINTALKEDAQVDLFRDITFYVYFS 180
Db 121 TFLQLERRKGVQSSGIMLTFWLVALVCALAILRSKINTALKEDAQVDLFRDITFYVYFS 180
QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPSSASFLSRTITFWITGLIVRGVQPLEGSD 240
Db 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPSSASFLSRTITFWITGLIVRGVQPLEGSD 240
QY 241 LWSLNKEDTSEQVVPVLVKNMKCEAKTRKQPKVYVSSKDPAPKPKSSKVDANEVEAL 300
Db 241 LWSLNKEDTSEQVVPVLVKNMKCEAKTRKQPKVYVSSKDPAPKPKSSKVDANEVEAL 300
QY 301 IVKSPQKEWNPSLFKVLKTYTGPFLMSFFFKAIHDLMMFSGPQLKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEWNPSLFKVLKTYTGPFLMSFFFKAIHDLMMFSGPQLKLLIKFVNDTKAPD 360
QY 361 WQGYFYTVLLFVTACLOTFLVHOFHICFVSGMRITKTAIVGAVYRKALVITNSARKSSTV 420
Db 361 WQGYFYTVLLFVTACLOTFLVHOFHICFVSGMRITKTAIVGAVYRKALVITNSARKSSTV 420

```



QY 421 GEIVNLSVDAORFMDLATYINNIWSAPLOVILALYLLWNLGSPSLAGVAVVWLVMPVN 480  
 Db 421 GEIVNLSVDAORFMDLATYINNIWSAPLOVILALYLLWNLGSPSLAGVAVVWLVMPVN 480  
 QY 481 AVAMKTKTYQVAHMKSKDNRIKIMNEIINGIKVKLKYAWELAFKDKVLAIQEEELVKLV 540  
 Db 481 AVAMKTKTYQVAHMKSKDNRIKIMNEIINGIKVKLKYAWELAFKDKVLAIQEEELVKLV 540  
 QY 541 KSNYLSAVCTFTWCTPPLVALCTFAVYTTIDENNILDAQTAFVSLALFNILRPNILP 600  
 Db 541 KSNYLSAVCTFTWCTPPLVALCTFAVYTTIDENNILDAQTAFVSLALFNILRPNILP 600  
 QY 601 MISSIVQASVSLKRLRIFLSHEELEPDSIERPVKGGTNSITVRNATFTWASDPT 660  
 Db 601 MISSIVQASVSLKRLRIFLSHEELEPDSIERPVKGGTNSITVRNATFTWASDPT 660  
 QY 661 LINGITSIPEGALVAVVGVQCGKSLLSALLAEMDKVEGHVAIKGSVAVVPQAWIQND 720  
 Db 661 LINGITSIPEGALVAVVGVQCGKSLLSALLAEMDKVEGHVAIKGSVAVVPQAWIQND 720  
 QY 721 SURENLLFCQLEEPYRVSIOACALLPDLLEILPSGDRTEIGEGKGNLGGOKQVSLAR 780  
 Db 721 SURENLLFCQLEEPYRVSIOACALLPDLLEILPSGDRTEIGEGKGNLGGOKQVSLAR 780  
 QY 781 AVYSNADIYLPDDPLSAVDAHVKGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 Db 781 AVYSNADIYLPDDPLSAVDAHVKGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 QY 841 MSGGKISEMGVQELLARGAFAEFLRTVASTEQDABENGVTGVSFGKEAKQWENG 900  
 Db 841 MSGGKISEMGVQELLARGAFAEFLRTVASTEQDABENGVTGVSFGKEAKQWENG 900  
 QY 901 LVTDSAGKLOQLSSSSSYSGDISRHNNSTAELOKAEAKKETWKLMEADKAQGVKL 960  
 Db 901 LVTDSAGKLOQLSSSSSYSGDISRHNNSTAELOKAEAKKETWKLMEADKAQGVKL 960  
 QY 961 SVYWDYMKAIGLFISFLIFPMCNHVSALAGNYWLSLWTDPIVNGTQEHKTVLSVYG 1020  
 Db 961 SVYWDYMKAIGLFISFLIFPMCNHVSALAGNYWLSLWTDPIVNGTQEHKTVLSVYG 1020  
 QY 1021 ALGISQGIATVGYSMVSTGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
 Db 1021 ALGISQGIATVGYSMVSTGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
 QY 1081 DTVDMSIPEVIMKFMGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140  
 Db 1081 DTVDMSIPEVIMKFMGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140  
 QY 1141 KXLESVRSPPVSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 Db 1141 KXLESVRSPPVSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 QY 1201 VRLCEVGNICIVLFAALFAVISHSISLGLVLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
 Db 1201 VRLCEVGNICIVLFAALFAVISHSISLGLVLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETAKEAPWQIETAPPSPWPQGVGRVERNYCLRYREDLDFVLRHINVTINGG 1320  
 Db 1261 VERLKEYSETAKEAPWQIETAPPSPWPQGVGRVERNYCLRYREDLDFVLRHINVTINGG 1320  
 QY 1321 EKVIGVGRGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 Db 1321 EKVIGVGRGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLRWNLDPFSQYDEEVTWLSLAHLKDFVSALPKDLHECARGENLSVGORQLVCL 1440  
 Db 1381 SGLRWNLDPFSQYDEEVTWLSLAHLKDFVSALPKDLHECARGENLSVGORQLVCL 1440  
 QY 1441 ABALLRKTILVDEATAVLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500  
 Db 1441 ABALLRKTILVDEATAVLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500  
 QY 1501 DKGEIQEYCAPSDLLQQRGLFYSMKADAGLV 1531

Db 1501 DKGEIQEYCAPSDLLQQRGLFYSMKADAGLV 1531  
 RESULT 11  
 US-08-463-179A-2  
 ; Sequence 2, Application US/08463179A  
 ; Patent No. 6001563  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cole, Susan P.C.  
 ; APPLICANT: Deeley, Roger G.  
 ; TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: LAHIVE & COCKFIELD  
 ; STREET: 60 State Street, suite 510  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/463,179A  
 ; FILING DATE:  
 ; CLASSIFICATION: 536  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/966,923  
 ; FILING DATE: 27-OCT-1992  
 ; APPLICATION NUMBER: 08/029,340  
 ; FILING DATE: 8-MAR-1993  
 ; APPLICATION NUMBER: 08/141,893  
 ; FILING DATE: 26-OCT-1993  
 ; APPLICATION NUMBER: 08/407,207  
 ; FILING DATE: 20-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: DeConti, Giulio A. Jr.  
 ; REGISTRATION NUMBER: 31,503  
 ; REFERENCE/DOCKET NUMBER: PQI-002CP8  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5941  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1531 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-463-179A-2  
 Query Match 78.4%; Score 7849; DB 3; Length 1531;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MALRGFCADGSDPLDWNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60  
 Db 1 MALRGFCADGSDPLDWNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFLYLSRH 60  
 QY 61 DRGIQMTPLNKTALGFLLIWVCWADLFVSPWERSRGIFLAPVFLVSPTLGTTLLA 120  
 Db 61 DRGIQMTPLNKTALGFLLIWVCWADLFVSPWERSRGIFLAPVFLVSPTLGTTLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVFS 180  
 Db 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVFS 180  
 QY 181 LLLIQLVLSCEFSRSPISFSETHDPNCPRESSASFLSRITFWWITGLIVRGYRQPLEGSD 240  
 Db 181 LLLIQLVLSCEFSRSPISFSETHDPNCPRESSASFLSRITFWWITGLIVRGYRQPLEGSD 240



QY 241 LMSLNKEDTSEQVPLVKNWKECAKTRKQPKVYVYSKDPAPKESKVDANEVEAL 300  
DB 241 LMSLNKEDTSEQVPLVKNWKECAKTRKQPKVYVYSKDPAPKESKVDANEVEAL 300  
QY 301 IVKSPQKWNPSLKVLYKTGPGPYFLMSPFFKAIHDLMMFSGPQILKLIKFVNDTKAPD 360  
DB 301 IVKSPQKWNPSLKVLYKTGPGPYFLMSPFFKAIHDLMMFSGPQILKLIKFVNDTKAPD 360  
QY 361 WQGYFYTVLLFVTTACLTQTLVHLQYFPHICFVSGMRKTAIVGAVYRKALVITNSARKSSTV 420  
DB 361 WQGYFYTVLLFVTTACLTQTLVHLQYFPHICFVSGMRKTAIVGAVYRKALVITNSARKSSTV 420  
QY 421 GEIVNLSVDQAFMDLATYINNIWSAPLOVILALYLWNLNLPSPVLGAVVWMLVMPVN 480  
DB 421 GEIVNLSVDQAFMDLATYINNIWSAPLOVILALYLWNLNLPSPVLGAVVWMLVMPVN 480  
QY 481 AVAMKTKTYQVAHMKSKDNRIKLNNEILNGIKVLKYAWELA FKDKVLAIROELKVLK 540  
DB 481 AVAMKTKTYQVAHMKSKDNRIKLNNEILNGIKVLKYAWELA FKDKVLAIROELKVLK 540  
QY 541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
DB 541 KSAYLSAVGTFTWCTPFLVALCTPAVYVTTIDENNILDAQTAFAVSLALFNILRFPNLILP 600  
QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKOGGTTNSITVRNATFTWASDPPT 660  
DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKOGGTTNSITVRNATFTWASDPPT 660  
QY 661 LINGITFISPEGALVAVGVGCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQWIOND 720  
DB 661 LINGITFISPEGALVAVGVGCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQWIOND 720  
QY 721 SURENIFCCOLEEPPYRSVIOACALLPDLLEILPSGDRTEIGEKGWNLSSGGQKQVSLAR 780  
DB 721 SURENIFCCOLEEPPYRSVIOACALLPDLLEILPSGDRTEIGEKGWNLSSGGQKQVSLAR 780  
QY 781 AVYSNADIYDFDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVLIIV 840  
DB 781 AVYSNADIYDFDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVLIIV 840  
QY 841 MSGGKISEMGSYOELLARDGAFELRTYASTEQDABEENGVTGVSFGPKAKOMENGM 900  
DB 841 MSGGKISEMGSYOELLARDGAFELRTYASTEQDABEENGVTGVSFGPKAKOMENGM 900  
QY 901 LVTDSAGKOLQRLSSSSSYSGDISRHNSSTAELKAEAKKEETWKLMEADKAQOGVKL 960  
DB 901 LVTDSAGKOLQRLSSSSSYSGDISRHNSSTAELKAEAKKEETWKLMEADKAQOGVKL 960  
QY 961 SVYWDYMKAIGLFISFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
DB 961 SVYWDYMKAIGLFISFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
QY 1021 ALGISQGTAVFGYSMAVSTGGILASRLHVDLHLSILSPMSFPFRTSGNLVNRFSKEL 1080  
DB 1021 ALGISQGTAVFGYSMAVSTGGILASRLHVDLHLSILSPMSFPFRTSGNLVNRFSKEL 1080  
QY 1081 DTVDMSIPBEVKMFWSLGNVIGACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140  
DB 1081 DTVDMSIPBEVKMFWSLGNVIGACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140  
QY 1141 KRLESVRSPPVSHFNETLLGVSVIRAFEEQERFIHQSDLVKVDENQKAYPSIVANRWLA 1200  
DB 1141 KRLESVRSPPVSHFNETLLGVSVIRAFEEQERFIHQSDLVKVDENQKAYPSIVANRWLA 1200  
QY 1261 VERLKEYSTEKEAPQIETAPPSPQVGVVERFNRYCLRYREDLDLFLVRHINTYINGG 1320  
DB 1261 VERLKEYSTEKEAPQIETAPPSPQVGVVERFNRYCLRYREDLDLFLVRHINTYINGG 1320  
QY 1321 EKVGVIGRTGAGKSSITLGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLV 1380

DB 1321 EKVGVIGRTGAGKSSITLGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLV 1380  
QY 1381 SSGSRMNLDPFSQYSDREBWTSLAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
DB 1381 SSGSRMNLDPFSQYSDREBWTSLAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
QY 1441 ARALLRKTILVDEATAAVALDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
DB 1441 ARALLRKTILVDEATAAVALDLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
QY 1501 DKGEIOEYGAPESDLLQORGLFYSMAKDAGLV 1531  
DB 1501 DKGEIOEYGAPESDLLQORGLFYSMAKDAGLV 1531

RESULT 12  
US-08-461-384B-2  
; Sequence 2, Application US/08461384B  
; Patent No. 6025473  
; GENERAL INFORMATION:  
; APPLICANT: Cole, Susan P.C.  
; APPLICANT: Deeley, Roger G.  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
; STREET: Queen's University at Kingston  
; CITY: Kingston  
; STATE: Ontario  
; COUNTRY: CANADA  
; ZIP: K7L 3N6  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,384B  
; FILING DATE: 05-JUN-95  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/966,923  
; FILING DATE: 27-OCT-1992  
; APPLICATION NUMBER: 08/029,340  
; FILING DATE: 8-MAR-1993  
; APPLICATION NUMBER: 08/141,893  
; FILING DATE: 26-OCT-1993  
; APPLICATION NUMBER: 08/407,207  
; FILING DATE: 20-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Steeg, Carol Miernicki  
; REGISTRATION NUMBER: 39,539  
; REFERENCE/DOCKET NUMBER: Q1547  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (613) 545-2342  
; TELEFAX: (613) 545-6853  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1531 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-461-384B-2

Query Match 78.4%; Score 7849; DB 3; Length 1531;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCGADGSDPLWDMNVTWNTSNPDFTKCFQNTLVWVPCFYLWACFPFYFLYLSRH 60  
DB 1 MALRGFCGADGSDPLWDMNVTWNTSNPDFTKCFQNTLVWVPCFYLWACFPFYFLYLSRH 60  
QY 61 DRGYIQMTPLNKTALGFLWIVCVADLFYSFWSRSRGIFLAPVFLVSPITLLGTTLLA 120

Db 61 DRGYIQMTPLNKTALGFLFWVCWADLFYFWEERSRGIFLAPVFLVSPITLIGITLLA 120  
 QY 121 TELIQLERRKGVOSSGIMLTFWLVVALCALAILRSKIMTALKEDAQVDLFRDITFVYFVS 180  
 Db 121 TELIQLERRKGVOSSGIMLTFWLVVALCALAILRSKIMTALKEDAQVDLFRDITFVYFVS 180  
 QY 181 LLLIQLVLCFSDRSPLFSETIHDNPNCPSSASFLSRITFWMTGLIIRGVRPLEGSD 240  
 Db 181 LLLIQLVLCFSDRSPLFSETIHDNPNCPSSASFLSRITFWMTGLIIRGVRPLEGSD 240  
 QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPKVYVSSKDPAPKPKSSKVDANEVEAL 300  
 Db 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPKVYVSSKDPAPKPKSSKVDANEVEAL 300  
 QY 301 IVKSPKQWNPFLFKVLYKTFPGYFLMSPFFKAHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 Db 301 IVKSPKQWNPFLFKVLYKTFPGYFLMSPFFKAHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFTVACLQTLVLHQQYFHCIFVSGMRKIKTAVIGAVYKALVITNSARKSTV 420  
 Db 361 WQGYFYTVLLFTVACLQTLVLHQQYFHCIFVSGMRKIKTAVIGAVYKALVITNSARKSTV 420  
 QY 421 GEIVNLMVSQAORFMDLATYINMWSAPLQVILALYLWNLGSPVLAGVAVWVLVMPVN 480  
 Db 421 GEIVNLMVSQAORFMDLATYINMWSAPLQVILALYLWNLGSPVLAGVAVWVLVMPVN 480  
 QY 481 AYWMKTKTYOVANHMSKDNRIKLMEILNGIKVLKYAWELAPKDKVLAIQBELKVLK 540  
 Db 481 AYWMKTKTYOVANHMSKDNRIKLMEILNGIKVLKYAWELAPKDKVLAIQBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFAVLSALFNILRFPNLILP 600  
 Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFAVLSALFNILRFPNLILP 600  
 QY 601 MVISSIVQASVSLKRLRPLSHEELEPDSIERRPVKDGCGTNSITVRNATFTWASDDPT 660  
 Db 601 MVISSIVQASVSLKRLRPLSHEELEPDSIERRPVKDGCGTNSITVRNATFTWASDDPT 660  
 QY 661 LNGITFTSPEGALVAVGVGCGKSLLSALLAENDKVEGHVAIKGSVAYVPOQAWIQND 720  
 Db 661 LNGITFTSPEGALVAVGVGCGKSLLSALLAENDKVEGHVAIKGSVAYVPOQAWIQND 720  
 QY 721 SLRENILFGCQLEEPYRSVIOACALLPDLPLPSGDRTEIGKGNVLSGGQKQVSLAR 780  
 Db 721 SLRENILFGCQLEEPYRSVIOACALLPDLPLPSGDRTEIGKGNVLSGGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840  
 Db 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840  
 QY 841 MSGGKISEMGVQOELLARDGAPAEFLRTYASTEQOEAENGVTGSGPGKEAKOMENG 900  
 Db 841 MSGGKISEMGVQOELLARDGAPAEFLRTYASTEQOEAENGVTGSGPGKEAKOMENG 900  
 QY 901 LVTDGAGLQORQLSSSSSYSGDISRHNSIAELQKAEAKKEETWKLMEADKAOTGOVKL 960  
 Db 901 LVTDGAGLQORQLSSSSSYSGDISRHNSIAELQKAEAKKEETWKLMEADKAOTGOVKL 960  
 QY 961 SVYWDYKAIGLFTSFLSIFLFCMCHVVSALASNYWLSLWTDPIVNGTOEHTKVLRSYVG 1020  
 Db 961 SVYWDYKAIGLFTSFLSIFLFCMCHVVSALASNYWLSLWTDPIVNGTOEHTKVLRSYVG 1020  
 QY 1021 ALGISQIAGVPGYSMAVSGIGLILASRCUHLVLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
 Db 1021 ALGISQIAGVPGYSMAVSGIGLILASRCUHLVLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
 QY 1081 DTVDMSIPEVIKMFMSLFNVIGACIVILLATPAAIIIPPLGLIYFFVQFYVASSRQL 1140  
 Db 1081 DTVDMSIPEVIKMFMSLFNVIGACIVILLATPAAIIIPPLGLIYFFVQFYVASSRQL 1140  
 QY 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200

Db 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLWLRMSSEMETNIVA 1260  
 Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLWLRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETKEAPWIOETAPPSSWPQVGRVEFRNYCLARYEDLDVLRHINTYINGG 1320  
 Db 1261 VERLKEYSETKEAPWIOETAPPSSWPQVGRVEFRNYCLARYEDLDVLRHINTYINGG 1320  
 QY 1321 EKVGVIGRTGACKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPDPPVLF 1380  
 Db 1321 EKVGVIGRTGACKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPDPPVLF 1380  
 QY 1381 SGLSRMNLDPFSQYSDSEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 Db 1381 SGLSRMNLDPFSQYSDSEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 QY 1441 ARALLURTKILVLOBEATAVDLETTDLLOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
 Db 1441 ARALLURTKILVLOBEATAVDLETTDLLOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
 QY 1501 DKGEIQEYGAQSDLLQORGLFYFYSMAKADAGLV 1531  
 Db 1501 DKGEIQEYGAQSDLLQORGLFYFYSMAKADAGLV 1531

RESULT 13

US-08-407-207A-2  
 ; Sequence 2, Application US/08407207A  
 ; Patent No. 6063621  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Deeley, Roger G.  
 ; APPLICANT: Cole, Susan P.C.  
 ; TITLE OF INVENTION: ANTIBODIES TO A MULTIDRUG RESISTANCE PROTEIN  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
 ; STREET: Queen's University at Kingston  
 ; CITY: Kingston  
 ; STATE: Ontario  
 ; COUNTRY: Canada  
 ; ZIP: K7L 3N6  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/407,207A  
 ; FILING DATE: 20-MAR-1995  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/966,923  
 ; FILING DATE: 27-OCT-1992  
 ; APPLICATION NUMBER: 08/029,340  
 ; FILING DATE: 8-MAR-1993  
 ; APPLICATION NUMBER: 08/141,893  
 ; FILING DATE: 26-OCT-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Steeg, Carol Miernicki  
 ; REGISTRATION NUMBER: 39,539  
 ; REFERENCE/DOCKET NUMBER: Q1512  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (613) 545-2342  
 ; TELEFAX: (613) 545-6853  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1531 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-407-207A-2

Query Match	78.4%;	Score 7849;	DB 3;	Length 1531;	
Best Local Similarity	99.9%;	Pred. No. 0;			
Matches 1529;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;	
QY	1	MALRFGCSADGSDPLWDMVNTWNTSNPDTKCFONTLVVWVFCFYLWACFPFYFLYLSRH	60		
DB	1	MALRFGCSADGSDPLWDMVNTWNTSNPDTKCFONTLVVWVFCFYLWACFPFYFLYLSRH	60		
QY	61	DRGYTQMTPLNKTALGELLWVCWADLFYFWERSRGIFLAPVPLVSPPTLLGTTLLA	120		
DB	61	DRGYTQMTPLNKTALGELLWVCWADLFYFWERSRGIFLAPVPLVSPPTLLGTTLLA	120		
QY	121	TELIQLERRKGVSQSGIMLTFMLVALVICALAILRSKIMTALKEDAQVDLFRDITFYVYFS	180		
DB	121	TELIQLERRKGVSQSGIMLTFMLVALVICALAILRSKIMTALKEDAQVDLFRDITFYVYFS	180		
QY	181	LLLIQLVLSCFSDRSPLETHDNPCESSASFLSRITFWITGLIIVRGYRQPLEGSD	240		
DB	181	LLLIQLVLSCFSDRSPLETHDNPCESSASFLSRITFWITGLIIVRGYRQPLEGSD	240		
QY	241	LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYVSSKDPAPKESKVDANEVEAL	300		
DB	241	LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYVSSKDPAPKESKVDANEVEAL	300		
QY	301	IVKSPQKEWNPFLKVLKTFPGPYFLMSFFFKAIHDLMMFSGPQILKLIKFVNDTKAPD	360		
DB	301	IVKSPQKEWNPFLKVLKTFPGPYFLMSFFFKAIHDLMMFSGPQILKLIKFVNDTKAPD	360		
QY	361	WQGYFTVLLFTVACLOTILVHQYHHI CFSVGMRIKTAIVIGAYRKALVITNSARKSTV	420		
DB	361	WQGYFTVLLFTVACLOTILVHQYHHI CFSVGMRIKTAIVIGAYRKALVITNSARKSTV	420		
QY	421	GEIVNLSMDAORFMDLATYINMISAPLOVILALYLLNLGSPSLAGVAVWMLMVPVN	480		
DB	421	GEIVNLSMDAORFMDLATYINMISAPLOVILALYLLNLGSPSLAGVAVWMLMVPVN	480		
QY	481	AVAMAKTKTYQVAHMKSKDNRIKLMNEILNGIKLVKLAYAWELAFKDKVLAIRQEBLVKL	540		
DB	481	AVAMAKTKTYQVAHMKSKDNRIKLMNEILNGIKLVKLAYAWELAFKDKVLAIRQEBLVKL	540		
QY	541	KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDQATFVSLALFNILRPFNLILP	600		
DB	541	KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDQATFVSLALFNILRPFNLILP	600		
QY	601	MVISSTVQASVSLKRLRIFLSHEELEPDSIERPVKGGTNSITVRNATFTWASDPPT	660		
DB	601	MVISSTVQASVSLKRLRIFLSHEELEPDSIERPVKGGTNSITVRNATFTWASDPPT	660		
QY	661	LANGITFSIPEGALVAVVGQVGGKSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND	720		
DB	661	LANGITFSIPEGALVAVVGQVGGKSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQND	720		
QY	721	SURENILFCQLEBPYRVSIVIOACALLPDLLEILPSGDRTEIGEKNVLSGGQKQVSLAR	780		
DB	721	SURENILFCQLEBPYRVSIVIOACALLPDLLEILPSGDRTEIGEKNVLSGGQKQVSLAR	780		
QY	781	AVYSNADIYLFDDPLSADVAHVGKHFENNIGPKMLKNKTRILVTHSMSYLPQVDVILV	840		
DB	781	AVYSNADIYLFDDPLSADVAHVGKHFENNIGPKMLKNKTRILVTHSMSYLPQVDVILV	840		
QY	841	MSGGKISEMGYSQELLARDGAEFLRTYASTEQEQAEEENGVTGVSFGPKAKQWENG	900		
DB	841	MSGGKISEMGYSQELLARDGAEFLRTYASTEQEQAEEENGVTGVSFGPKAKQWENG	900		
QY	901	LVTDSAGLQQLQSSSSSYSGDISRHNSSTAELOKAEAKBETWKLMEADKAQTQGVKL	960		
DB	901	LVTDSAGLQQLQSSSSSYSGDISRHNSSTAELOKAEAKBETWKLMEADKAQTQGVKL	960		
QY	961	SVYWDYMKAGLGFISFLSIFLFCMNHVSALSNYLSLWTDPIVNGTOEHTKVLRSVYG	1020		
DB	961	SVYWDYMKAGLGFISFLSIFLFCMNHVSALSNYLSLWTDPIVNGTOEHTKVLRSVYG	1020		

RESULT 14

US-08-463-092B-6  
 ; Sequence 6, Application US/08463092B  
 ; Patent No. 5766880  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cole, Susan P.C.  
 ; APPLICANT: Delevy, Roger G.  
 ; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING  
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS  
 ; STREET: Queen's University at Kingston  
 ; CITY: Kingston  
 ; STATE: Ontario  
 ; COUNTRY: CANADA  
 ; ZIP: K7L 3N6  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/463,092B  
 ; FILING DATE: 05-JUN-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/966,923  
 ; FILING DATE: 27-OCT-1992  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/029,340  
 ; FILING DATE: 8-MAR-1993  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/141,893

QY	1021	ALGISQGIAGVFGYSMAVSIIGGILASRCCLHVDLLHLSILRSPMFFERTPSGNNLVNRFSKEL	1080		
DB	1021	ALGISQGIAGVFGYSMAVSIIGGILASRCCLHVDLLHLSILRSPMFFERTPSGNNLVNRFSKEL	1080		
QY	1081	DTVDWSMIEPVTKMFNGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSROL	1140		
DB	1081	DTVDWSMIEPVTKMFNGSLFNIVGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSROL	1140		
QY	1141	KELSESRSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA	1200		
DB	1141	KELSESRSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA	1200		
QY	1201	VRLECVGNCIVLFAALFAVISRHSLSAGLVLSVSYQVTTYLLNWLVRMSSEMETNIVA	1260		
DB	1201	VRLECVGNCIVLFAALFAVISRHSLSAGLVLSVSYQVTTYLLNWLVRMSSEMETNIVA	1260		
QY	1261	VERLKEYSETKEAPWQIQETAPPSSWPQVGRVFNRYCLRVREDLDVFLRHINVTINGG	1320		
DB	1261	VERLKEYSETKEAPWQIQETAPPSSWPQVGRVFNRYCLRVREDLDVFLRHINVTINGG	1320		
QY	1321	EKVGIVGRTGAGKSSLTGLPRINESARGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380		
DB	1321	EKVGIVGRTGAGKSSLTGLPRINESARGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF	1380		
QY	1381	SGSLRMNLDPPFSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL	1440		
DB	1381	SGSLRMNLDPPFSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL	1440		
QY	1441	ARALARKTKILVLDATAVDDTLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL	1500		
DB	1441	ARALARKTKILVLDATAVDDTLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL	1500		
QY	1501	DKGEIQEYCAPSDLLQQRGLFYSMADAGLV	1531		
DB	1501	DKGEIQEYCAPSDLLQQRGLFYSMADAGLV	1531		

```

; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1546
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1528 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-0928-6

Query Match          69.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGFCSDGSDPLMDMNVNTWNTSNPDTKCFQNTVLVWPCFYLWACFPFYLRLSRH 60
DB 1 MALRFSACSADGSDPLMDMNVNTWNTSNPDTKCFQNTVLVWPCFYLWACFPFYLRLSRH 60
QY 61 DRGYQMTPLNTKTKALGFLLMVVCWADLFYFWSRSGIFLAPVFLVSPITLLGITLLA 120
DB 61 DRGYQMTPLNTKTKALGFLLMVVCWADLFYFWSRSGIFLAPVFLVSPITLLGITLLA 120
QY 121 TELIOLERRKGVOSSIMLTFWLVALCALATLRSKIMTALKEDAQVDLFRDITFYVFS 180
DB 121 TELIOLERRKGVOSSIMLTFWLVALCALATLRSKIMTALKEDAQVDLFRDITFYVFS 180
QY 181 LLLIQLVLSGSDRPLSETHDPNCPPESSASFLSRITFWITGLIVRGVRQPLEGSD 240
DB 181 LVLVLQVLSCFSDCPLSETHDPNCPPESSASFLSRITFWITGLIVRGVRQPLEGSD 240
QY 241 LWSLNKEDTSEOVVPLVKNWKECAKTRKQPKVYVS-SKDPAPQKSSSKVDANEEVEA 299
DB 241 LWSLNKEDTSEOVVPLVKNWKECAKTRKQPKVYVS-SKDPAPQKSSSKVDANEEVEA 300
QY 300 LIVKSPQKNEPSELFKLYKTFPGPYFLMSFPFKAHDMFSGPOILKLLIKFVNDTKAP 359
DB 301 LIVKSPHDKREPSLFKLYKTFPGPYFLMSFLYKALHDMFAGPKILELIINFVNDREAP 360
QY 360 DWQGYFTYVTLFVTAQLTLVHLQYFHICFVSGMRITKAVIGAVYRKALVITNARKSST 419
DB 361 DWQGYFTYVTLFVTAQLTLVHLQYFHICFVSGMRITKAVIGAVYRKALVITNARKSST 420
QY 420 VGEIYNLMSVDAQRPMDLATYINMWSAPLOVILALYLLNLGSPVLAVGVMLVMPV 479
DB 421 VGEIYNLMSVDAQRPMDLATYINMWSAPLOVILALYLLNLGSPVLAVGVMLVMPV 480
QY 480 NAVMAKTKTYQVAHMKSDNRKILMNEILNGIKVLKYAWELAFKDKVLAIROBELKVL 539
DB 481 NAVMAKTKTYQVAHMKSDNRKILMNEILNGIKVLKYAWELAFKDKVLAIROBELKVL 540
QY 540 KGSAYLSAVGFTTWCTPFLVALCTFAVYVITDENNILDQAFAVSLALFNILAFPLN 599
DB 541 KGSAYLSAVGFTTWCTPFLVALCTFAVYVITDENNILDQAFAVSLALFNILAFPLN 600
QY 600 PWLTISSIVQASVSLKRLIFLSHELEPDSTERRPVKGGTNSITVNNATFTWARSDDP 659
DB 601 PWLTISSIVQASVSLKRLIFLSHELEPDSTERRPVKGGTNSITVNNATFTWARSDDP 659
QY 660 TLNGITFSPGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAKGSVAVVPOQAWION 719
DB 660 TLNGITFSPGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAKGSVAVVPOQAWION 719

```

```

QY 720 DSLRENILFGCOLLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKGWNLGGQKORVSLA 779
DB 720 DSLRENILFGHPLQENYKAVMEACALLPDLLEILPSGDRTEIGEKGWNLGGQKORVSLA 779
QY 780 RAVYSNADIIYLFDDPLSAVDHVGKHIFENVIKPGMLKNKTRILVTHSMVSLPQVDVII 839
DB 780 RAVYSNADIIYLFDDPLSAVDHVGKHIFEKVVGPMGLLNKTRILVTHGSIPLQVDVII 839
QY 840 VMSGKISSEMGYSQELLARDGAFAPFLRTYASTEQEODAEENGVTGVSQPGKEAKOMENG 899
DB 840 VMSGKISSEMGYSQELLARDGAFAPFLRTYASTEQEODAEENGVTGVSQPGKEAKOMENG 896
QY 900 MLVTDGAKQORQLSSSSSYSGDISRHNNHTAELOKAEAKKEETWKLMEADKATGVQK 959
DB 900 MLVTDGAKQORQLSSSSSYSGDISRHNNHTAELOKAEAKKEETWKLMEADKATGVQK 955
QY 960 LSVYWDYMKALGFLFISLIFLFCNHYVSLASALNYWLSLWTDLPIVNGTOEHTKVRSLV 1018
DB 960 LSVYWDYMKALGFLFISLIFLFCNHYVSLASALNYWLSLWTDLPIVNGTOEHTKVRSLV 1015
QY 1019 YGALGISQIAVFGYSMAVSGTILASCLHVDLLHLSILRSPMSFFERTPSGNLVNRFPSK 1078
DB 1016 YGALGILQGAIFGYSMAVSGTILASCLHVDLLHLSILRSPMSFFERTPSGNLVNRFPSK 1075
QY 1079 ELDTVDSMIPVIVKMGSLFNVIAGACIVILLATPIAAIIIPPLGLIYVFFQRFVASSR 1138
DB 1076 ELDTVDSMIPVIVKMGSLFNVIAGACIVILLATPIAAIIIPPLGLIYVFFQRFVASSR 1135
QY 1139 QLKELSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRW 1198
DB 1136 QLKELSVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRW 1195
QY 1199 LAVLECVGNCIVLPAALPAVISHSLAGLVGSVSLQVTVTYLNLVMSSEMETNI 1258
DB 1196 LAVLECVGNCIVLPAALPAVISHSLAGLVGSVSLQVTVTYLNLVMSSEMETNI 1255
QY 1259 VAVERLKEYSTKEAPWQIQTAPPSPQVQVGFERNYCLRYREDLDLVLKHNVTIN 1318
DB 1256 VAVERLKEYSTKEAPWQIQTAPPSPQVQVGFERNYCLRYREDLDLVLKHNVTIN 1315
QY 1319 GGEKVGIVGRTGAGKSSITLGLFRINESAGEEIIIDGINIAKIGLHDLRFKITTIPQDPV 1378
DB 1316 GGEKVGIVGRTGAGKSSITLGLFRINESAGEEIIIDGINIAKIGLHDLRFKITTIPQDPV 1375
QY 1379 LFSGLSNLMDPPSOYSDDEEVTSLAHLKDFVSALPDKLDHCEACGGENLSVGORQLV 1438
DB 1376 LFSGLSNLMDPPSOYSDDEEVTSLAHLKDFVSALPDKLDHCEACGGENLSVGORQLV 1435
QY 1439 CLARALLRKTILVLDATAVLDLTDLLIQTSTRTQFEDCTVLTIAHRLNTIMDYTRVI 1498
DB 1436 CLARALLRKTILVLDATAVLDLTDLLIQTSTRTQFEDCTVLTIAHRLNTIMDYTRVI 1495
QY 1499 VLDKGTIOEYCAPSLLDLOQRGLFYSMAKADAGLV 1531
DB 1496 VLDKGTIOEYCAPSLLDLOQRGLFYSMAKADAGLV 1528

RESULT 15
US-08-462-109A-6
; Sequence 6, Application US/08462109A
; Patent No. 5882875
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P. C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA

```

```

; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,109A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: PQI-002CP4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1528 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-462-109A-6

```

```

Query Match      69.9%; Score 7002.5; DB 2; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGCSDGSDPLDNDVNTWNTNSPDFTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60
DB 1 MALRGCSDGSDPLDNDVNTWNTNSPDFTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60
QY 61 DRGYQMTPLNKTALGELLWVWADLFYFWSERSRGIFLAPVFLVSPITLLGITLLA 120
DB 61 DRGYQMTPLNKTALGELLWVWADLFYFWSERSRGIFLAPVFLVSPITLLGITLLA 120
QY 121 TELIOLERRKGVOSSGIMTLFWLVALVCAIILRSKIMTALKEDAQVDLFRDITFYVPS 180
DB 121 TELIOLERRKGVOSSGIMTLFWLVALVCAIILRSKIMTALKEDAQVDLFRDITFYVPS 180
QY 181 LLIQIOLVLSCFSDRPLRSETTHDPNCPSSASFLSRITFWITGLIYRGYRQPLEGSD 240
DB 181 LLIQIOLVLSCFSDRPLRSETTHDPNCPSSASFLSRITFWITGLIYRGYRQPLEGSD 240
QY 241 LWSLNKEDTSEQVVPVLVNNWKEKCAKTRKQPVVYIS-SKDPAPQKSSKYDANEVEEA 299
DB 241 LWSLNKEDTSEQVVPVLVNNWKEKCAKTRKQPVVYIS-SKDPAPQKSSKYDANEVEEA 299
QY 300 LIVKSPQKWNPSLFVLYKTPGPFYLMSSFFKAIHDLMMFSGPQILKLLIKFVNDTKAP 359
DB 300 LIVKSPQKWNPSLFVLYKTPGPFYLMSSFFKAIHDLMMFSGPQILKLLIKFVNDTKAP 359
QY 360 DMQGYFYTVLLVFTACLOTLVHQPFCFVSGMRITKTAIVGAVYRKALVINTSARKSST 419
DB 360 DMQGYFYTVLLVFTACLOTLVHQPFCFVSGMRITKTAIVGAVYRKALVINTSARKSST 419
QY 420 VGEIVNLSVDAQRFDLTYINMWSAPLQVILALYLWNLGSPVLAGVAVMLVMPV 479
DB 420 VGEIVNLSVDAQRFDLTYINMWSAPLQVILALYLWNLGSPVLAGVAVMLVMPV 479
QY 480 NAVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLATROBELKVL 539
DB 480 NAVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLATROBELKVL 539
QY 481 NAVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFQDKVMSIQBELKVL 540
DB 481 NAVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFQDKVMSIQBELKVL 540

```

Search completed: March 18, 2005, 11:03:47  
Job time : 48.7091 secs

```

QY 540 KKSAYLSAVGTFWCTPFLVALCTFAVYVTTIDENNILDAQTAFVSLALFNILRPLNIL 599
DB 540 KKSAYLSAVGTFWCTPFLVALCTFAVYVTTIDENNILDAQTAFVSLALFNILRPLNIL 599
QY 600 PMWISSIVQASVSLKRLRIFLSHELEPDSITERPVKGGGTNSITVNAFTTWARSOPP 659
DB 600 PMWISSIVQASVSLKRLRIFLSHELEPDSITERPVKGGGTNSITVNAFTTWARSOPP 659
QY 660 TLNGITTSIPGALVAVVGVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQWQ 719
DB 660 TLNGITTSIPGALVAVVGVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVYVQQAQWQ 719
QY 720 DSLRENILFGCOLPEPYRSVIOCALLPDLEILPSGDRTEIGEKGVLNLSGQKQVSLA 779
DB 720 DSLRENILFGCOLPEPYRSVIOCALLPDLEILPSGDRTEIGEKGVLNLSGQKQVSLA 779
QY 780 RAVYSNADIYLFDDPLSAVDHVGKHIFENVIKPKMLKNTKTRILVTHSMVLPQVDVII 839
DB 780 RAVYSNADIYLFDDPLSAVDHVGKHIFENVIKPKMLKNTKTRILVTHSMVLPQVDVII 839
QY 840 VMGGKISEMGSYOELLARDGAFAEFLRTYASTEOEODAEENGVTGVSQPKGKAQWENG 899
DB 840 VMGGKISEMGSYOELLARDGAFAEFLRTYASTEOEODAEENGVTGVSQPKGKAQWENG 899
QY 900 MLVTDVGVKHLQRLHLSNSSSHSGDTSQOHSIAELQKAGA-KEETWKLMEADKAGTQVQ 959
DB 900 MLVTDVGVKHLQRLHLSNSSSHSGDTSQOHSIAELQKAGA-KEETWKLMEADKAGTQVQ 959
QY 960 LSVTWDYKKAIGLFIPLFMCNVHVSALASNYWLSLWTD--PIVNGTQEHKTVRLSV 1018
DB 960 LSVTWDYKKAIGLFIPLFMCNVHVSALASNYWLSLWTD--PIVNGTQEHKTVRLSV 1018
QY 1019 YGALGISOGIAVFGYSMAVSIIGGILASRLCLHVDLLHSILRGPMSPFERTPSGNLVNRSK 1078
DB 1019 YGALGISOGIAVFGYSMAVSIIGGILASRLCLHVDLLHSILRGPMSPFERTPSGNLVNRSK 1078
QY 1079 ELDTVDVSMIPEVIMKPMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVASSR 1138
DB 1079 ELDTVDVSMIPEVIMKPMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFVASSR 1138
QY 1139 QKRLSVRSRVPVSHFNETLLGVSVIRAFEOERFIHQSDLKVDENKAYVPSIVANRW 1198
DB 1139 QKRLSVRSRVPVSHFNETLLGVSVIRAFEOERFIHQSDLKVDENKAYVPSIVANRW 1198
QY 1199 LAVRLECVGNCIVLFAALFAVIRSHLSAGLVGLSVSYSLQVITAYLNLVMSSEMETNI 1258
DB 1199 LAVRLECVGNCIVLFAALFAVIRSHLSAGLVGLSVSYSLQVITAYLNLVMSSEMETNI 1258
QY 1259 VAVRELKEYSETEKEAPWQIOETAPPSPQVGRVFRNYCLRYREDLDFVLRHINVTIN 1318
DB 1259 VAVRELKEYSETEKEAPWQIOETAPPSPQVGRVFRNYCLRYREDLDFVLRHINVTIN 1318
QY 1319 GGEKGVIGRTGAGKSSITGLIFRINESAGEIIIDGINIAKIGHDLRFXKTIIPQDPV 1378
DB 1319 GGEKGVIGRTGAGKSSITGLIFRINESAGEIIIDGINIAKIGHDLRFXKTIIPQDPV 1378
QY 1379 LFGSGLRNLDPPFSQYSDVEEYVMTSLAHLKDFVSALPKDHECAEGENLSVQORQV 1438
DB 1379 LFGSGLRNLDPPFSQYSDVEEYVMTSLAHLKDFVSALPKDHECAEGENLSVQORQV 1438
QY 1439 CLARALLRKTILVLDEATAAVDLETDLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVI 1498
DB 1439 CLARALLRKTILVLDEATAAVDLETDLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVI 1498
QY 1499 VLDKGEIYEGAPSDLLQORGLFYSMKADAGLV 1531
DB 1499 VLDKGEIYEGAPSDLLQORGLFYSMKADAGLV 1531
QY 1531 VLDKGEIYEGAPSDLLQORGLFYSMKADAGLV 1531
DB 1531 VLDKGEIYEGAPSDLLQORGLFYSMKADAGLV 1531

```

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 11:03:55 ; Search time 615.662 Seconds  
(without alignments)  
1045.320 Million cell updates/sec

Title: US-10-665-283-8  
Perfect score: 10016  
Sequence: 1 MALRGFCSADGSDPLWDNV.....RSVAVAKPKFSISPDLS 1947

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1401741 seqs, 330541175 residues  
Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/ECT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/ECTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	78.5	1531	11	US-09-939-853A-86
2	7860	78.5	1531	16	US-10-667-891-6
3	7860	78.5	1531	17	US-10-484-577-678
4	7769	77.6	1515	11	US-09-939-853A-87
5	7037	70.3	1388	16	US-10-408-765A-1718
6	7002.5	69.9	1528	11	US-09-939-853A-88
7	4491.5	44.8	1527	11	US-09-939-853A-83
8	4491.5	44.8	1527	11	US-09-939-853A-84
9	4491.5	44.8	1527	15	US-10-295-027-1342
10	4396	43.9	1522	11	US-09-939-853A-85
11	3860.5	38.5	1548	16	US-10-667-891-1
12	3603.5	36.0	1505	16	US-10-667-891-4
13	3601.5	36.0	1499	16	US-10-667-891-3

14	3600	35.9	1542	15	US-10-363-112-4	Sequence 4, Appli
15	3600	35.9	1545	15	US-10-252-155-4	Sequence 4, Appli
16	3600	35.9	1545	15	US-10-252-155-24	Sequence 24, Appli
17	3600	35.9	1545	15	US-10-252-155-26	Sequence 26, Appli
18	3600	35.9	1545	15	US-10-252-155-28	Sequence 28, Appli
19	3600	35.9	1545	15	US-10-252-155-30	Sequence 30, Appli
20	3600	35.9	1545	15	US-10-252-155-32	Sequence 32, Appli
21	3600	35.9	1545	15	US-10-252-155-34	Sequence 34, Appli
22	3600	35.9	1545	15	US-10-252-155-38	Sequence 38, Appli
23	3600	35.9	1545	15	US-10-252-155-40	Sequence 40, Appli
24	3600	35.9	1545	15	US-10-252-155-46	Sequence 46, Appli
25	3600	35.9	1545	15	US-10-252-155-48	Sequence 48, Appli
26	3600	35.9	1545	15	US-10-363-112-2	Sequence 2, Appli
27	3600	35.9	1545	15	US-10-363-112-6	Sequence 6, Appli
28	3600	35.9	1545	15	US-10-363-112-8	Sequence 8, Appli
29	3600	35.9	1545	15	US-10-363-112-10	Sequence 10, Appli
30	3600	35.9	1545	15	US-10-363-112-12	Sequence 12, Appli
31	3600	35.9	1545	15	US-10-363-112-14	Sequence 14, Appli
32	3600	35.9	1545	15	US-10-363-112-16	Sequence 16, Appli
33	3599.5	35.9	1507	16	US-10-667-891-5	Sequence 5, Appli
34	3599	35.9	1545	15	US-10-252-155-42	Sequence 42, Appli
35	3597	35.9	1545	15	US-10-252-155-44	Sequence 44, Appli
36	3595	35.9	1545	16	US-10-741-601-369	Sequence 369, App
37	3595	35.9	1545	16	US-10-741-601-370	Sequence 370, App
38	3593	35.9	1545	15	US-10-252-155-36	Sequence 36, Appli
39	3582	35.8	1545	15	US-10-252-155-604	Sequence 604, App
40	3385.5	33.8	1503	10	US-09-792-616-3	Sequence 3, Appli
41	3385.5	33.8	1503	16	US-10-764-328-3	Sequence 3, Appli
42	3385.5	33.8	1503	17	US-10-741-600-1583	Sequence 1583, Ap
43	3289	32.8	1494	15	US-10-369-493-6697	Sequence 6697, Ap
44	3225	32.2	1573	15	US-10-369-493-6867	Sequence 6867, Ap
45	3225	32.2	1573	15	US-10-369-493-6868	Sequence 6868, Ap

ALIGNMENTS

RESULT 1  
US-09-939-853A-86  
; Sequence 86, Application US/09939853A  
; Publication No. US20040039163A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess et al.  
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-099  
; CURRENT APPLICATION NUMBER: US/09/939,853A  
; CURRENT FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: 60/228,191  
; PRIOR FILING DATE: 2000-08-25  
; PRIOR APPLICATION NUMBER: 60/267,300  
; PRIOR FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: 60/269,961  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/277,337  
; PRIOR FILING DATE: 2001-03-20  
; NUMBER OF SEQ ID NOS: 159  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 86  
; LENGTH: 1531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-939-853A-86

Query Match	78.5%	Score 7860;	DB 11;	Length 1531;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 1531;	Conservative	0;		Gaps 0;
Qy	1	MALRGFCSADGSDPLWDNVNTWNTSNPDFTKCFQNTLVWVPCFYLVWVPCFYLYLSRH	60	
Db	1	MALRGFCSADGSDPLWDNVNTWNTSNPDFTKCFQNTLVWVPCFYLVWVPCFYLYLSRH	60	
Qy	61	DRGIQMTPLNKTALGLFWLWVWADLFVSFWERSRGIFLAPVFLVSPILLGLTTLLA	120	

Db 61 DRGYIQTMTPLNKTALGFLMWIVCWADLPYSFWERSRGIFLAPVFLVPTLLGITTLA 120  
QY 121 TFLIQLERRKGVOSSGIMLTFWLVALVCAIALIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
Db 121 TFLIQLERRKGVOSSGIMLTFWLVALVCAIALIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPPCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
Db 181 LLLIQLVLSCFSDRSPLFSETIHDNPPCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
QY 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPKVYVSSKDPAPKPKSSKVDANEVEAL 300  
Db 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPKVYVSSKDPAPKPKSSKVDANEVEAL 300  
QY 301 IVKSPKQWNPSLFLKYLKTFGDFLMSFFFAIHDLMMFSPQIILKLIKFVNDTKAPD 360  
Db 301 IVKSPKQWNPSLFLKYLKTFGDFLMSFFFAIHDLMMFSPQIILKLIKFVNDTKAPD 360  
QY 361 WQGYFTVLLFVTACLTQLVHLQYPHI CFVSGWRIKTA VIGAVYRKALVITNSARKSSTV 420  
Db 361 WQGYFTVLLFVTACLTQLVHLQYPHI CFVSGWRIKTA VIGAVYRKALVITNSARKSSTV 420  
QY 421 GEIVNLSVDAQRFMDLATYININMSAPLOVILALYLLNLGSPSLAGVAVVWLMVPVN 480  
Db 421 GEIVNLSVDAQRFMDLATYININMSAPLOVILALYLLNLGSPSLAGVAVVWLMVPVN 480  
QY 481 AVWAMKTKYQVAHMKSKDNRIKLNIEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
Db 481 AVWAMKTKYQVAHMKSKDNRIKLNIEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
QY 541 KSAYLSAVGTFWVCTPFLVALCTFAVYVTTIDENNILDAQTAFVLSLALFNILRFPNLILP 600  
Db 541 KSAYLSAVGTFWVCTPFLVALCTFAVYVTTIDENNILDAQTAFVLSLALFNILRFPNLILP 600  
QY 601 MWISSIVQASVSLKRLIRIFLSHEELEPDSIERPVPKGGTINSITVRNATFTWASDPPT 660  
Db 601 MWISSIVQASVSLKRLIRIFLSHEELEPDSIERPVPKGGTINSITVRNATFTWASDPPT 660  
QY 661 LNGITFISPEGALVAVGVGGCKSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIOWD 720  
Db 661 LNGITFISPEGALVAVGVGGCKSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIOWD 720  
QY 721 SURENIFGCOLLEPYRYSVIOACALLPDLPLPSGDRTEIGEKGWNLGGOKQVSLAR 780  
Db 721 SURENIFGCOLLEPYRYSVIOACALLPDLPLPSGDRTEIGEKGWNLGGOKQVSLAR 780  
QY 781 AVYSNADIYLPDPLSADVAHVCKHIFENVIGPKGMLKNKTRILVTHSMYLPQVDVIV 840  
Db 781 AVYSNADIYLPDPLSADVAHVCKHIFENVIGPKGMLKNKTRILVTHSMYLPQVDVIV 840  
QY 841 MSGGKISEMGYSQELLARDGAPAEFLRTVASTEQDQAEENGVTGSGPGKEAKQWENG 900  
Db 841 MSGGKISEMGYSQELLARDGAPAEFLRTVASTEQDQAEENGVTGSGPGKEAKQWENG 900  
QY 901 LVTDAGKOLQRLSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQTQVKL 960  
Db 901 LVTDAGKOLQRLSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQTQVKL 960  
QY 961 SVYDYMKAIGLIFSLFSLFPMCNHVSALASNYWLSLWTDPIVNGQEHKVRLSVYG 1020  
Db 961 SVYDYMKAIGLIFSLFSLFPMCNHVSALASNYWLSLWTDPIVNGQEHKVRLSVYG 1020  
QY 1021 ALGISQGIAGVGYMAVSGIGTILASRCLHVDLLHSILRSPMFPERTPSGNLVNRFSEL 1080  
Db 1021 ALGISQGIAGVGYMAVSGIGTILASRCLHVDLLHSILRSPMFPERTPSGNLVNRFSEL 1080  
QY 1081 DTVDNMIPEVIMKFMGSLFNIVIGACIVILLATPAAIIIPPLGLIYFPVQFYVASSRQL 1140  
Db 1081 DTVDNMIPEVIMKFMGSLFNIVIGACIVILLATPAAIIIPPLGLIYFPVQFYVASSRQL 1140  
QY 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFTHQSDLVKVDENOKAYPSIVANRWLA 1200  
Db 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFTHQSDLVKVDENOKAYPSIVANRWLA 1200

QY 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVGLSVSVLSQVTTYLNLWLRMSSEMETNIYA 1260  
Db 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVGLSVSVLSQVTTYLNLWLRMSSEMETNIYA 1260  
QY 1261 VERLKEYSETEKEAPWQIQTETAPPSWPQVGRVEFRNCLRYREDLDFVLRHINTYINGG 1320  
Db 1261 VERLKEYSETEKEAPWQIQTETAPPSWPQVGRVEFRNCLRYREDLDFVLRHINTYINGG 1320  
QY 1321 EKVGIVGRTGAGKSSLTGLCPRIINESARGEIIIDGINIAKIGHDLRPFKTIIPDPVLF 1380  
Db 1321 EKVGIVGRTGAGKSSLTGLCPRIINESARGEIIIDGINIAKIGHDLRPFKTIIPDPVLF 1380  
QY 1381 SSSLRMNLDPPSQSDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
Db 1381 SSSLRMNLDPPSQSDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
QY 1441 ARALLRKTILVLDDEATAVDLETTDLTQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Db 1441 ARALLRKTILVLDDEATAVDLETTDLTQSTIRTFQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
QY 1501 DKGEIQEYCAPSDLLIQRGLFYSMADAGLV 1531  
Db 1501 DKGEIQEYCAPSDLLIQRGLFYSMADAGLV 1531

## RESULT 2

US-10-667-891-6  
; Sequence 6, Application US/10667891  
; Publication No. US20040171024A1  
; GENERAL INFORMATION:  
; APPLICANT: ROTH, CHARLES W.  
; APPLICANT: BREY, PAUL T.  
; APPLICANT: HOLM, INGE  
; APPLICANT: GRAILLES, MARINE  
; APPLICANT: RZHETSKY, ANDREY  
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND  
; TITLE OF INVENTION: ANOPHELES  
; FILE REFERENCE: 03495-0294-00000  
; CURRENT APPLICATION NUMBER: US/10/667,891  
; CURRENT FILING DATE: 2003-09-23  
; PRIOR APPLICATION NUMBER: 60/413,469  
; PRIOR FILING DATE: 2002-09-26  
; NUMBER OF SEQ ID NOS: 76  
; SOFTWARE: Patent in Ver. 3.2  
; SEQ ID NO 6  
; LENGTH: 1531  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-667-891-6

Query Match 78.5%; Score 7860; DB 16; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDWNTWNTSNDFTKCFONTVLVWVPCFYLWACFPFYLISRH 60  
Db 1 MALRGFCSADGSDPLMDWNTWNTSNDFTKCFONTVLVWVPCFYLWACFPFYLISRH 60  
QY 61 DRGYIQTMTPLNKTALGFLMWIVCWADLPYSFWERSRGIFLAPVFLVSPPTLLGITTLA 120  
Db 61 DRGYIQTMTPLNKTALGFLMWIVCWADLPYSFWERSRGIFLAPVFLVSPPTLLGITTLA 120  
QY 121 TFLIQLERRKGVOSSGIMLTFWLVALVCAIALIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
Db 121 TFLIQLERRKGVOSSGIMLTFWLVALVCAIALIRSKIMTALKEDAQVDLFRDITFYVYFS 180  
QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPPCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
Db 181 LLLIQLVLSCFSDRSPLFSETIHDNPPCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
QY 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPKVYVSSKDPAPKPKSSKVDANEVEAL 300  
Db 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPKVYVSSKDPAPKPKSSKVDANEVEAL 300



Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPKESKVDANEVEAL 300  
 QY 301 IVKSPKKNWPSLFKVLKTPGYPFLMSFFFAIHDLMFSGPQILKLLIKFVNDTKAPD 360  
 Db 301 IVKSPKKNWPSLFKVLKTPGYPFLMSFFFAIHDLMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTTACQLTLVHLQYFHI CFVSGMRKTA VIGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTTACQLTLVHLQYFHI CFVSGMRKTA VIGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLWNLGSPSLAGVAVMLVMPVN 480  
 Db 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLWNLGSPSLAGVAVMLVMPVN 480  
 QY 481 AVAMKTKTYQVAHMKSKDNRIKLNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540  
 Db 481 AVAMKTKTYQVAHMKSKDNRIKLNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFVSLALFNILRFFLNILP 600  
 Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFVSLALFNILRFFLNILP 600  
 QY 601 MVISSIVQASVLSKLRIFLSHEELEPDSIERPPVKDGGTNSITVRNATFTWASDDPT 660  
 Db 601 MVISSIVQASVLSKLRIFLSHEELEPDSIERPPVKDGGTNSITVRNATFTWASDDPT 660  
 QY 661 LINGITFSIPEGALVAVVGQVCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIOND 720  
 Db 661 LINGITFSIPEGALVAVVGQVCGKSLLSALLAEMDKVEGHVAIKGSVAVYVQQAWIOND 720  
 QY 721 SURENILFCQLEEPYRSVIOACALLPDLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
 Db 721 SURENILFCQLEEPYRSVIOACALLPDLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
 QY 781 AVYSNADIYFDPLDSAVDAHVGHKIFENVIGPKMLKNKTRILVTHSMSYLPQVDVLIIV 840  
 Db 781 AVYSNADIYFDPLDSAVDAHVGHKIFENVIGPKMLKNKTRILVTHSMSYLPQVDVLIIV 840  
 QY 841 MSGGKISEMGSTQELLARDAFAELRTYASTEQEODAEENGVTGSGPKAKOMENGM 900  
 Db 841 MSGGKISEMGSTQELLARDAFAELRTYASTEQEODAEENGVTGSGPKAKOMENGM 900  
 QY 901 LVTDSAGKQLQRLSSSSSYSGDI SRHNSSTAELKAEAKKEETWKLMEADKAQTVGVKL 960  
 Db 901 LVTDSAGKQLQRLSSSSSYSGDI SRHNSSTAELKAEAKKEETWKLMEADKAQTVGVKL 960  
 QY 961 SVYDYMKAIGLFI SFLSIFLFCMCHVSA LAGNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 Db 961 SVYDYMKAIGLFI SFLSIFLFCMCHVSA LAGNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 QY 1021 ALGISOGIAVFGYSMAVSI GGLIARCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEL 1080  
 Db 1021 ALGISOGIAVFGYSMAVSI GGLIARCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEL 1080  
 QY 1081 DTVDMSIMEVIMKMGSLFNVI GACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140  
 Db 1081 DTVDMSIMEVIMKMGSLFNVI GACIVILLATPIAAIIPPLGLIYFFVQRFYVASSRQL 1140  
 QY 1141 KRLESVRSRSPVSHFNETLLGVSVIRAPFEQERFIHQSDLKVDENQKAYIPSIIVANRWLA 1200  
 Db 1141 KRLESVRSRSPVSHFNETLLGVSVIRAPFEQERFIHQSDLKVDENQKAYIPSIIVANRWLA 1200  
 QY 1201 VRLECVGNCI VLPFAALFAVISRHSLSAGLVGSVSLQVTTYLWNLVRMSSEMETNIVA 1260  
 Db 1201 VRLECVGNCI VLPFAALFAVISRHSLSAGLVGSVSLQVTTYLWNLVRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKAPWQIQTAPPSSWPQVGRVEFNRYCLRYREDLDLFLRHINTVINGG 1320  
 Db 1261 VERLKEYSETEKAPWQIQTAPPSSWPQVGRVEFNRYCLRYREDLDLFLRHINTVINGG 1320  
 QY 1321 EKVIGVGRGTAGKSSLTGLFRINSAEGEIIIDGINIAKIGHDLRFKTIIPQDPVLIF 1380  
 Db 1321 EKVIGVGRGTAGKSSLTGLFRINSAEGEIIIDGINIAKIGHDLRFKTIIPQDPVLIF 1380

QY 1381 SGSLRMNLDPPFSQXSDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSYGQRLVCL 1440  
 Db 1381 SGSLRMNLDPPFSQXSDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSYGQRLVCL 1440  
 QY 1441 ARALLRKTKILVDEBATAAVIDLETDLLIQTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTKILVDEBATAAVIDLETDLLIQTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIOEYGA PSDLLQORGLFYSMADAGLV 1531  
 Db 1501 DKGEIOEYGA PSDLLQORGLFYSMADAGLV 1531

RESULT 3

US-10-484-577-678  
 ; Sequence 678, Application US/10484577  
 ; Publication No. US20050032724A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: EPIDAUROS Biotechnologie Aktiengesellschaft  
 ; TITLE OF INVENTION: Means and methods for improved treatment of cancer based on UGRI  
 ; FILE REFERENCE: F2285PCT-1  
 ; CURRENT APPLICATION NUMBER: US/10/484,577  
 ; CURRENT FILING DATE: 2004-01-22  
 ; PRIOR APPLICATION NUMBER: PCT/EP 02/08220  
 ; PRIOR FILING DATE: 2002-07-23  
 ; PRIOR APPLICATION NUMBER: EP 01 11 7608.8  
 ; PRIOR FILING DATE: 2001-07-23  
 ; PRIOR APPLICATION NUMBER: EP 02011710.7  
 ; PRIOR FILING DATE: 2002-05-24  
 ; NUMBER OF SEQ ID NOS: 683  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 678  
 ; LENGTH: 1531  
 ; TYPE: PR1  
 ; ORGANISM: Homo sapiens  
 ; US-10-484-577-678

Query Match 78.5%; Score 7860; DB 17; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSDGSDPLWDMNVTWNTSNPDTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60  
 Db 1 MALRGFCSDGSDPLWDMNVTWNTSNPDTKCFQNTVLVWVPCFYLWACFPFYLYLSRH 60  
 QY 61 DRGYIQMTPLNKTALGFLWIVCWADLFYSFWERSGIFLAPVFLVSPILLGTTLLA 120  
 Db 61 DRGYIQMTPLNKTALGFLWIVCWADLFYSFWERSGIFLAPVFLVSPILLGTTLLA 120  
 QY 121 TFLIOLERRKGQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFVYVFS 180  
 Db 121 TFLIOLERRKGQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFVYVFS 180  
 QY 181 LLLIQLVLSGSDRSPLSFSETHDNPCCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
 Db 181 LLLIQLVLSGSDRSPLSFSETHDNPCCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPKESKVDANEVEAL 300  
 Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPKESKVDANEVEAL 300  
 QY 301 IVKSPKKNWPSLFKVLKTPGYPFLMSFFFAIHDLMFSGPQILKLLIKFVNDTKAPD 360  
 Db 301 IVKSPKKNWPSLFKVLKTPGYPFLMSFFFAIHDLMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTTACQLTLVHLQYFHI CFVSGMRKTA VIGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTTACQLTLVHLQYFHI CFVSGMRKTA VIGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLWNLGSPSLAGVAVMLVMPVN 480  
 Db 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYLWNLGSPSLAGVAVMLVMPVN 480

```

QY 481 AVAMKTKTYQVAHMSKONRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQBELKVLK 540
Db 481 AVAMKTKTYQVAHMSKONRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQBELKVLK 540
QY 541 KSAYLSAVGFTTWCPTFELVALCTFAVYVTDENNILDQAQAFVSLALFNILRFPNINLP 600
Db 541 KSAYLSAVGFTTWCPTFELVALCTFAVYVTDENNILDQAQAFVSLALFNILRFPNINLP 600
QY 601 MWISSIVQASVSLKRLIRIFLSHELEPDSIERRPVKDGCGTNSITVRNATFTWASDPT 660
Db 601 MWISSIVQASVSLKRLIRIFLSHELEPDSIERRPVKDGCGTNSITVRNATFTWASDPT 660
QY 661 LAGITPSIEGALVAVVGVCGCKSLLSALLAEMDKVGHVAIKGSVAVVPCQAWIWD 720
Db 661 LAGITPSIEGALVAVVGVCGCKSLLSALLAEMDKVGHVAIKGSVAVVPCQAWIWD 720
QY 721 SURENIFGQLEPYRYSVIOACALLPDLPLPSGDRTEIGEGVNLGGQKQVSLAR 780
Db 721 SURENIFGQLEPYRYSVIOACALLPDLPLPSGDRTEIGEGVNLGGQKQVSLAR 780
QY 781 AVYSNADIYLFDDPLSADAHGKHIFENVIKPKMLKNKTRILVTHSMSYLPQVDVIIV 840
Db 781 AVYSNADIYLFDDPLSADAHGKHIFENVIKPKMLKNKTRILVTHSMSYLPQVDVIIV 840
QY 841 MSGGKISEMSGYQELLARCAFAELRTYASTEQEOADAENGVTGSGPKAKOMENGM 900
Db 841 MSGGKISEMSGYQELLARCAFAELRTYASTEQEOADAENGVTGSGPKAKOMENGM 900
QY 901 LVTDSAGKOLQOLSSSSSYSGDISRHNSSTAELKAEAKKEETWKLMEADKAQGVKL 960
Db 901 LVTDSAGKOLQOLSSSSSYSGDISRHNSSTAELKAEAKKEETWKLMEADKAQGVKL 960
QY 961 SYVDYMKAGLFIPLSIFLPMCHNVSAALSNYLSLWTDPIVNGTQEHKTVLSYVG 1020
Db 961 SYVDYMKAGLFIPLSIFLPMCHNVSAALSNYLSLWTDPIVNGTQEHKTVLSYVG 1020
QY 1021 ALGISQGIAGVGYSMASVIGGILASRCLHVDLLHSILSPMSFFERTPSGNLVNRFSEL 1080
Db 1021 ALGISQGIAGVGYSMASVIGGILASRCLHVDLLHSILSPMSFFERTPSGNLVNRFSEL 1080
QY 1081 DTVDSMIEVVKMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140
Db 1081 DTVDSMIEVVKMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRYVASSRQL 1140
QY 1141 KELESVSRSPVYSHNEILLGVSVIRAFEEQBRFHQSDLVKDNOKAYPSIVANRWLA 1200
Db 1141 KELESVSRSPVYSHNEILLGVSVIRAFEEQBRFHQSDLVKDNOKAYPSIVANRWLA 1200
QY 1201 VRLECVGNCIVLFAALFAVIRHSLASAGLVGLSVSYSLQVTTYLNWLVRMSSEMETNIVA 1260
Db 1201 VRLECVGNCIVLFAALFAVIRHSLASAGLVGLSVSYSLQVTTYLNWLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETKEAPWQIQTAPSSWPQGVREFRNYCLRYREDLDVPLRHINVTINGG 1320
Db 1261 VERLKEYSETKEAPWQIQTAPSSWPQGVREFRNYCLRYREDLDVPLRHINVTINGG 1320
QY 1321 EKVGVGRTGAGKSLTLGLFRINESARGEIIIDGINIAKIGLHDLRPKTIIPQDPVLF 1380
Db 1321 EKVGVGRTGAGKSLTLGLFRINESARGEIIIDGINIAKIGLHDLRPKTIIPQDPVLF 1380
QY 1381 SGSLRMNLDPFSQYSDEEVTSLAHLKDFVSALPKDLHECAEGGNSLVGQRLVCL 1440
Db 1381 SGSLRMNLDPFSQYSDEEVTSLAHLKDFVSALPKDLHECAEGGNSLVGQRLVCL 1440
QY 1441 ARALLRKTIILVLEATAVDLETDLLIQSTIRTPQEDCTVLTIAHRLNTIMDYTRIVL 1500
Db 1441 ARALLRKTIILVLEATAVDLETDLLIQSTIRTPQEDCTVLTIAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYCAPSLLQQRGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEYCAPSLLQQRGLFYSMAKDAGLV 1531

```

```

RESULT 4
US-09-939-853A-87
; Sequence 87, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgees et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 1515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-87

```

```

Query Match 77.6%; Score 7769; DB 11; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 DWNVTWNTSNDDPKFCQNTVLVWVPCFYLWACPFYFLYLSRDRGHIQMTPLNKTKTA 76
Db 17 DWNVTWNTSNDDPKFCQNTVLVWVPCFYLWACPFYFLYLSRDRGHIQMTPLNKTKTA 76
QY 77 LGFLLWVWADLVFVSWERSRGIFLAPVFLVSPFLTGITTTLLATFLIQLERRGVQSSG 136
Db 77 LGFLLWVWADLVFVSWERSRGIFLAPVFLVSPFLTGITTTLLATFLIQLERRGVQSSG 136
QY 61 LGFLLWVWADLVFVSWERSRGIFLAPVFLVSPFLTGITTTLLATFLIQLERRGVQSSG 120
Db 61 LGFLLWVWADLVFVSWERSRGIFLAPVFLVSPFLTGITTTLLATFLIQLERRGVQSSG 120
QY 137 IMLTFWLVVALCALAILRSKIMTKALKEADAQVDLFRDITFVYVYFSLLLIQLVLSCFSDRSP 196
Db 137 IMLTFWLVVALCALAILRSKIMTKALKEADAQVDLFRDITFVYVYFSLLLIQLVLSCFSDRSP 196
QY 197 LFSFTIHDNPPCPSSASFSLRITFWITGLIVRGYQPLEGSLWSINKEDTSEQVVPV 256
Db 197 LFSFTIHDNPPCPSSASFSLRITFWITGLIVRGYQPLEGSLWSINKEDTSEQVVPV 256
QY 181 LFSFTIHDNPPCPSSASFSLRITFWITGLIVRGYQPLEGSLWSINKEDTSEQVVPV 240
Db 181 LFSFTIHDNPPCPSSASFSLRITFWITGLIVRGYQPLEGSLWSINKEDTSEQVVPV 240
QY 257 LVKNWKECAKTRKQPVKVYVSSKDPAPKPESSKVDANEVEALIVKSPQKWNPSLPKV 316
Db 257 LVKNWKECAKTRKQPVKVYVSSKDPAPKPESSKVDANEVEALIVKSPQKWNPSLPKV 316
QY 317 LYKTFGYPFLMSFPFKAIHDLMMFSGPQIILKLLIKFVNDTKAPDWQGYFYTVLLFVTACL 376
Db 317 LYKTFGYPFLMSFPFKAIHDLMMFSGPQIILKLLIKFVNDTKAPDWQGYFYTVLLFVTACL 376
QY 301 LYKTFGYPFLMSFPFKAIHDLMMFSGPQIILKLLIKFVNDTKAPDWQGYFYTVLLFVTACL 360
Db 301 LYKTFGYPFLMSFPFKAIHDLMMFSGPQIILKLLIKFVNDTKAPDWQGYFYTVLLFVTACL 360
QY 377 QTLVLHGYFHICFVSGMIRITAVIGAVYRKALVITNSARKSSTVGEIVNLSVDAQRFMD 436
Db 377 QTLVLHGYFHICFVSGMIRITAVIGAVYRKALVITNSARKSSTVGEIVNLSVDAQRFMD 436
QY 361 QTLVLHGYFHICFVSGMIRITAVIGAVYRKALVITNSARKSSTVGEIVNLSVDAQRFMD 420
Db 361 QTLVLHGYFHICFVSGMIRITAVIGAVYRKALVITNSARKSSTVGEIVNLSVDAQRFMD 420
QY 437 LATYINMWSAPLQVTLALYLLMLNLGSPSLAGVAVMVLMPVNVNMAKTKTYQVAHMK 496
Db 437 LATYINMWSAPLQVTLALYLLMLNLGSPSLAGVAVMVLMPVNVNMAKTKTYQVAHMK 496
QY 497 SKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQBELKVLKKSAYLSAVGFTTWC 556
Db 497 SKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQBELKVLKKSAYLSAVGFTTWC 556
QY 481 SKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQBELKVLKKSAYLSAVGFTTWC 540
Db 481 SKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQBELKVLKKSAYLSAVGFTTWC 540
QY 557 PFLVALCTFAVYVTDENNILDQAQAFVSLALFNILRFPNINLPMTSSIVQASVSLKRL 616
Db 557 PFLVALCTFAVYVTDENNILDQAQAFVSLALFNILRFPNINLPMTSSIVQASVSLKRL 616
QY 541 PFLVALCTFAVYVTDENNILDQAQAFVSLALFNILRFPNINLPMTSSIVQASVSLKRL 600
Db 541 PFLVALCTFAVYVTDENNILDQAQAFVSLALFNILRFPNINLPMTSSIVQASVSLKRL 600
QY 617 RIFLSHEELEBPSDIERRPVKDGCGTNSITVRNATFTWASDPTLNGITFISPEGLAVV 676
Db 617 RIFLSHEELEBPSDIERRPVKDGCGTNSITVRNATFTWASDPTLNGITFISPEGLAVV 676

```

```

Db 601 RIFLSHELEPDSIERRPVKGGGNSITVRNATFTWARSDDPTLNGITFISPEGALVAV 660
Qy 677 VQVCGCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIOWDSIRENILFGCOLEBPY 736
Db 661 VQVCGCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIOWDSIRENILFGCOLEBPY 720
Qy 737 VRSVTOACALLPDLEILPSGRTEIGEKNVLSGGQKORVSLARAVYSNADIYLPDDPLS 796
Db 721 VRSVTOACALLPDLEILPSGRTEIGEKNVLSGGQKORVSLARAVYSNADIYLPDDPLS 780
Qy 797 AVDAHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVLIIVMSGGKISMGSYOELL 856
Db 781 AVDAHVGKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVLIIVMSGGKISMGSYOELL 840
Qy 857 ARDGAFAFLRYTASTEOQDAEENGVTGVSFGKEAKOMENGLVTDTSAGKQLORQLSS 916
Db 841 ARDGAFAFLRYTASTEOQDAEENGVTGVSFGKEAKOMENGLVTDTSAGKQLORQLSS 900
Qy 917 SSSYSGDISRHNSIAELQAKAEKBEWKLMEADKAOTGQVKLSVYWDYMKAI GLFISF 976
Db 901 SSSYSGDISRHNSIAELQAKAEKBEWKLMEADKAOTGQVKLSVYWDYMKAI GLFISF 960
Qy 977 LSIFLFCMCHVSALASNYWLSLWTDPIVNGTQETHKVLRSVYGALGISQGI AVEGYSWA 1036
Db 961 LSIFLFCMCHVSALASNYWLSLWTDPIVNGTQETHKVLRSVYGALGISQGI AVEGYSWA 1020
Qy 1037 VSIGGILASRCLHVDLHLSILASPMSPFERTPSGNLVRFSKELDTVDSMIPEVIKMPWG 1096
Db 1021 VSIGGILASRCLHVDLHLSILASPMSPFERTPSGNLVRFSKELDTVDSMIPEVIKMPWG 1080
Qy 1097 SLFNVIGACIVTLLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPVYSHFN 1156
Db 1081 SLFNVIGACIVTLLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPVYSHFN 1140
Qy 1157 ETLGVSIVRAPEEGERFHQSDIKVDENOKAYPSIVANRWLA VRLSCVNCIVLFAAL 1216
Db 1141 ETLGVSIVRAPEEGERFHQSDIKVDENOKAYPSIVANRWLA VRLSCVNCIVLFAAL 1200
Qy 1217 FAVISRHLSL SAGLVLSYSIQVTTYLNWLRMSSEMETNIVAVERLKEYSETEKEAPW 1276
Db 1201 FAVISRHLSL SAGLVLSYSIQVTTYLNWLRMSSEMETNIVAVERLKEYSETEKEAPW 1260
Qy 1277 QIQTAPPSSWPQGRVFRNRYCLAYREDLDFVLRHINVTNGKBKGVIGVGTGAGKSSL 1336
Db 1261 QIQTAPPSSWPQGRVFRNRYCLAYREDLDFVLRHINVTNGKBKGVIGVGTGAGKSSL 1320
Qy 1337 TLGLFRINESAGEIIIDGINIAKIGLHDLREKTIIPQDPVL FSGSLRMNLDPPSOYSD 1396
Db 1321 TLGLFRINESAGEIIIDGINIAKIGLHDLREKTIIPQDPVL FSGSLRMNLDPPSOYSD 1380
Qy 1397 EEWMTSLELAHLKDFVSALPKDLDECAEGGENLSVGORQLVCLARALLRKTILVLDBA 1456
Db 1381 EEWMTSLELAHLKDFVSALPKDLDECAEGGENLSVGORQLVCLARALLRKTILVLDBA 1440
Qy 1457 TAAVDLETDLLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIQEYGAPSDLLQ 1516
Db 1441 TAAVDLETDLLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIQEYGAPSDLLQ 1500
Qy 1517 ORGLFYSMAKDAGLV 1531
Db 1501 ORGLFYSMAKDAGLV 1515

```

RESULT 5

```

US-10-408-765A-1718
; Sequence 1718, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.

```

```

; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 2077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1718
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-408-765A-1718

Query Match 70.3%; Score 7037; DB 16; Length 1388;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1387; Conservative 0; Mismatches 1; Indels 42; Gaps 1;

Qy 67 MTPLNKTALGFLLIWIVCWADLPYSFWERSRGIFLAPVFLVSPPTLLGITTLLATFLIQL 126
Db 1 MTPLNKTALGFLLIWIVCWADLPYSFWERSRGIFLAPVFLVSPPTLLGITTLLATFLIQL 60
Qy 127 ERRKGVSQSGIMLTFWLVALYCALAILRSKIMTALKEDAQVDLFRDITFYVYFSLLLQL 186
Db 61 ERRKGVSQSGIMLTFWLVALYCALAILRSKIMTALKEDAQVDLFRDITFYVYFSLLLQL 97
Qy 187 VLSCFSDRSPLFSETHDPNCPSSASFLSRIITFWMITGLIVRGYROPLEGSDLSLNK 246
Db 98 -----NPCSSASFLSRIITFWMITGLIVRGYROPLEGSDLSLNK 138
Qy 247 EDTSEQVVPVLVQNMKECAKTRQPKVYVSSKDPAPQKESKVDANEEVEALIVKSPQ 306
Db 139 EDTSEQVVPVLVQNMKECAKTRQPKVYVSSKDPAPQKESKVDANEEVEALIVKSPQ 198
Qy 307 KEWNPSPKVLVYKTFPGPYFLMSPPFKAIHDLMMSPGPOILKLIKFNVDTKAPDQOGPY 366
Db 199 KEWNPSPKVLVYKTFPGPYFLMSPPFKAIHDLMMSPGPOILKLIKFNVDTKAPDQOGPY 258
Qy 367 TVLLFVTAACLTQLVLHOFVHICFVSGMRIKTAIVGAVYRKALVITNSARKSSTVGEIVNL 426
Db 259 TVLLFVTAACLTQLVLHOFVHICFVSGMRIKTAIVGAVYRKALVITNSARKSSTVGEIVNL 318
Qy 427 MSVDAQRFMDLATYINMIWSAPLQVILALYLLNLGSPVLGAVVMVLPVNAVMAK 486
Db 319 MSVDAQRFMDLATYINMIWSAPLQVILALYLLNLGSPVLGAVVMVLPVNAVMAK 378
Qy 487 TKTQVAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKDKVLATROBELKVLKKSAYLS 546
Db 379 TKTQVAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKDKVLATROBELKVLKKSAYLS 438
Qy 547 AVGTFTVCTPFLVALCTFAVVTIDENNILDAOTAFVSLALFNILRPLNLPVVISI 606
Db 439 AVGTFTVCTPFLVALCTFAVVTIDENNILDAOTAFVSLALFNILRPLNLPVVISI 498
Qy 607 VQASVSLKRLRIFLSHEELEPDSIERREPVDKGGTNSITVRNATFTWARSDDPTLNGITF 666
Db 499 VQASVSLKRLRIFLSHEELEPDSIERREPVDKGGTNSITVRNATFTWARSDDPTLNGITF 558
Qy 667 SIPEGALVAVVGVQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIOWDSIRENI 726
Db 559 SIPEGALVAVVGVQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIOWDSIRENI 618
Qy 727 LFGCOLEBPYVRSVTOACALLPDLEILPSGRTEIGEKNVLSGGQKORVSLARAVYNA 786
Db 619 LFGCOLEBPYVRSVTOACALLPDLEILPSGRTEIGEKNVLSGGQKORVSLARAVYNA 678
Qy 787 DIYLFDDPLSADVAHVGHKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVLIIVMSGGKI 846
Db 679 DIYLFDDPLSADVAHVGHKHI FENVIGPKGMLKNKTRILVTHSMSYLPQVDVLIIVMSGGKI 738
Qy 847 SEMGSYOBELLARDGAFELRTYASTEOQDAEENGVTGVSFGKEAKOMENGLVTDTSAG 906

```

```

Db 739 SEMGYSQELLARDGAFAEFLRTYASTEQBDAAENGVTGVSFGKQAKQWENGLVTDGA 798
Qy 907 GKOLQROLSSSSYSDDISRHNSSTAELQAKAKKEETWKLMEADKAQTQGVKLSYWDY 966
Db 799 GKOLQROLSSSSYSDDISRHNSSTAELQAKAKKEETWKLMEADKAQTQGVKLSYWDY 858
Qy 967 MKAIGLFIISFLSIFPMCNHVSALASNYWLSLWTDPIVNGTQEHKVKRLSVYGALGISQ 1026
Db 859 MKAIGLFIISFLSIFPMCNHVSALASNYWLSLWTDPIVNGTQEHKVKRLSVYGALGISQ 918
Qy 1027 GIAVFGYSMAVSTGGILASRCLHVDLLHSILSPMSFFERTPSGNLVNRFSEKELDTVDSM 1086
Db 919 GIAVFGYSMAVSTGGILASRCLHVDLLHSILSPMSFFERTPSGNLVNRFSEKELDTVDSM 978
Qy 1087 IPEVIMKMGSLFNIVGACIVILLATPIAAIIIPGLGLIYFFVQRFYVASSQLKRLSEV 1146
Db 979 IPEVIMKMGSLFNIVGACIVILLATPIAAIIIPGLGLIYFFVQRFYVASSQLKRLSEV 1038
Qy 1147 SRSPVYSHNETLLGVSIVRAFEQEERFIHQSDLVKVDENOKAYPSIVANRWLAVRLECV 1206
Db 1039 SRSPVYSHNETLLGVSIVRAFEQEERFIHQSDLVKVDENOKAYPSIVANRWLAVRLECV 1098
Qy 1207 GNCIVLFAALFAVIRSHSLASAGLVLSYSYSQVTTYLNWLVRMSSEMETNIVAVERLKE 1266
Db 1099 GNCIVLFAALFAVIRSHSLASAGLVLSYSYSQVTTYLNWLVRMSSEMETNIVAVERLKE 1158
Qy 1267 YSETEKEAPWQIOETAPSSWPQVGRVFRNRYCLAYREDLDVLIHINVTINGGKVGIV 1326
Db 1159 YSETEKEAPWQIOETAPSSWPQVGRVFRNRYCLAYREDLDVLIHINVTINGGKVGIV 1218
Qy 1327 GRTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRPKTIIPQDPVLFSGSLRM 1386
Db 1219 GRTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRPKTIIPQDPVLFSGSLRM 1278
Qy 1387 NLDPSQYSDEBWTSLSLAHUKDFVSALPKLDHECAEGGNSLVGQRLVCLARALLR 1446
Db 1279 NLDPSQYSDEBWTSLSLAHUKDFVSALPKLDHECAEGGNSLVGQRLVCLARALLR 1338
Qy 1447 KTKILVLDEATAVLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTR 1496
Db 1339 KTKILVLDEATAVLETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTR 1388

RESULT 6
US-09-939-853A-88
; Sequence 88, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 88
; LENGTH: 1528
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-88

```

Query Match 69.9%; Score 7002.5; DB 11; Length 1528;  
 Best Local Similarity 88.0%; Pred. No. 0;  
 Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

```

Qy 1 MALRFGCSADGSDPLMDNNTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFYLSLH 60
Db 1 MALRFGCSADGSDPLMDNNTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFYLSLH 60
Qy 61 DRGYTQMTPLNKTALGFLWICWADLFYSFWERSRGIFLAPVLPVSPTLGLITLLA 120
Db 61 DRGYTQMTPLNKTALGFLWICWADLFYSFWERSRGIFLAPVLPVSPTLGLITLLA 120
Qy 121 TFLIOLRRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAOVDFRDIIFYVPS 180
Db 121 TFLIOLRRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAOVDFRDIIFYVPS 180
Qy 181 LLLIOLVLSGSDRSPLESEITHDNPCPSSASFLSRITFWIITGLIVRGVROPLESD 240
Db 181 LLLIOLVLSGSDRSPLESEITHDNPCPSSASFLSRITFWIITGLIVRGVROPLESD 240
Qy 241 LWSLNKEDTSEVVPVLVNNKKECDKSRKQPVRIYAPPKDPKPKSSQLDVNEEVEA 299
Db 241 LWSLNKEDTSEVVPVLVNNKKECDKSRKQPVRIYAPPKDPKPKSSQLDVNEEVEA 299
Qy 300 LIVKSPQKEMNPSLFKVLKYTFGPFYFMSFPFKAHIDLMFSGPOILKLLIKFVNDTKAP 359
Db 300 LIVKSPQKEMNPSLFKVLKYTFGPFYFMSFPFKAHIDLMFSGPOILKLLIKFVNDTKAP 359
Qy 360 DWQGVYFTVLLFVTAQTLVHGYFHTCFVSGWRIKTAVIGAVYRKALVITNSARKSST 419
Db 360 DWQGVYFTVLLFVTAQTLVHGYFHTCFVSGWRIKTAVIGAVYRKALVITNSARKSST 419
Qy 420 VGEIVNLSVDAQRFMDIATYINMIWSAPLOVILALYLLNLNLGSPVLAGVAVMLVMPV 479
Db 420 VGEIVNLSVDAQRFMDIATYINMIWSAPLOVILALYLLNLNLGSPVLAGVAVMLVMPV 479
Qy 480 NAVMAKTKTVQVAHMKSKONRIKMLNEILNGIKVLKIYAWELAFKDKVLAIRSEELKVL 539
Db 480 NAVMAKTKTVQVAHMKSKONRIKMLNEILNGIKVLKIYAWELAFKDKVLAIRSEELKVL 539
Qy 540 KKSAYLSAVGFTWVCTPPLVALCTFAVYVTIDENILDAOTAFVSLALFNILRPLNIL 599
Db 540 KKSAYLSAVGFTWVCTPPLVALCTFAVYVTIDENILDAOTAFVSLALFNILRPLNIL 599
Qy 600 PMWISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGCGGTNSITVRNATFTTWARSDPP 659
Db 600 PMWISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGCGGTNSITVRNATFTTWARSDPP 659
Qy 660 TNGITFSPREGALVAVGVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWTQN 719
Db 660 TNGITFSPREGALVAVGVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWTQN 719
Qy 720 DSLRENILFGQLEPPYRSVIOACALLPDLLEILPSGDRTEIGKGVNLSGGQKORVSLA 779
Db 720 DSLRENILFGQLEPPYRSVIOACALLPDLLEILPSGDRTEIGKGVNLSGGQKORVSLA 779
Qy 780 RAVYSNADIYLFDDPLSAVDHVGKHFENYVIGPKGMLKNKTRILVTHSMYSILPQVDVII 839
Db 780 RAVYSNADIYLFDDPLSAVDHVGKHFENYVIGPKGMLKNKTRILVTHSMYSILPQVDVII 839
Qy 840 VMSGGKISEMGSYQELLARDGAFAEFLRTYASTEQBDAAENGVTGVSFGKQAKQWENGL 896
Db 840 VMSGGKISEMGSYQELLARDGAFAEFLRTYASTEQBDAAENGVTGVSFGKQAKQWENGL 896
Qy 900 MLVTDGAKQIQROLSSSSYSDDISRHNSSTAELQAKAKKEETWKLMEADKAQTQGVK 959
Db 900 MLVTDGAKQIQROLSSSSYSDDISRHNSSTAELQAKAKKEETWKLMEADKAQTQGVK 959
Qy 959 LSVYWNWYKAIGLFIISFLSIFPMCNHVSALASNYWLSLWTDPIVNGTQEHKVKRLSV 1018
Db 959 LSVYWNWYKAIGLFIISFLSIFPMCNHVSALASNYWLSLWTDPIVNGTQEHKVKRLSV 1018
Qy 1019 YGALIGISOGIATVFGYSMAVSTGGILASRCLHVDLLHSILSPMSFFERTPSGNLVNRFSE 1078
Db 1019 YGALIGISOGIATVFGYSMAVSTGGILASRCLHVDLLHSILSPMSFFERTPSGNLVNRFSE 1078
Qy 1078 ELDTVDSMIEPIKMFMSGLFNIVGACIVILLATPIAAIIIPGLGLIYFFVQRFYVASSR 1138

```

Db	1076	ELDTVDNMIPQVIMKFMGSLFSVGAVIIILLATPIAAVIPPGLGVYFFVQFYVASSR	1135
Qy	1139	QKRLGESVSRSPVYSHFNETLLGVSVIRAFEEQBRFIHQSDLKVDENOKAYPSIVANRW	1198
Db	1136	QKRLGESVSRSPVYSHFNETLLGVSVIRAFEEQBRFIHQSDLKVDENOKAYPSIVANRW	1195
Qy	1199	LAVRLECGNCGIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLWRMSEMETNI	1258
Db	1196	LAVRLECGNCGIVLFAALFAVISRHSLSAGLVGLSVSYSLQITAYLNLWRMSEMETNI	1255
Qy	1259	VAVERLKEYSETKEAPWQIQETAPPSWPQVGRVEPRNTCLRYREDLDLDFVLKHINVTIN	1318
Db	1256	VAVERLKEYSETKEAPWQIQETAPPSWPQVGRVEPRNTCLRYREDLDLDFVLKHINVTIE	1315
Qy	1319	GGEKVGIVGRTGAGKSSLTGLFRINSAEGEIIIDGINIAKTGLHDLRKFITIIIPQDPV	1378
Db	1316	GGEKVGIVGRTGAGKSSLTGLFRINSAEGEIIIDGVNIAKTGLHNLRFKITIIIPQDPV	1375
Qy	1379	LFSGSLRNLNDPPFSQYDDEEVTLSLELAHLKDFVSALPDKLDHECAGGENLSVGORQLV	1438
Db	1376	LFSGSLRNLNDPPFSQYDDEEVTWMALELAHLKGFVSALPDKNHECAGGENLSVGORQLV	1435
Qy	1439	CLARALLRKTILVLDATAAVDLETDDLLQSTIRTOFEDCTVLITIAHRLNTIMDYTRVI	1498
Db	1436	CLARALLRKTILVLDATAAVDLETDDLLQSTIRTOFEDCTVLITIAHRLNTIMDYTRVI	1495
Qy	1499	VLDKGETQYQCAPSDDLQORGLFYFSMAKDAGLV	1531
Db	1496	VLDKGEYRECAPSELLOORGIFYFSMAKDAGLV	1528

## RESULT 7

```

US-09-939-853A-83
; Sequence 83, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/259,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 83
; LENGTH: 1527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-83

```

	Query Match	44.8%;	Score 4491.5;	DB 11;	Length 1527;
	Best Local Similarity	56.7%;	Pred. No. 0;		
	Matches	873;	Conservative 260;	Mismatches 371;	Indels 35; Gaps 9;
Qy	8	SADGSDPLMDNWTNTSNEDFKCFONTVLVWVPCFYLWACPPFFLYLSRHRDGYIQM	67		
Db	7	SGEIGSKFWDNSLVSHTYENPDLTPCFQNSLLAWVPCTIYLWVALLFCVLLYLRRHCRGYIIL	66		
Qy	68	TPLNKTKTALGFLMIWCWADLPFSYFWSRSGIFLAPVFLVSPSTLCITTLTLLATFLIQLE	127		
Db	67	SHLSKLRWLVGLVLLNCVSWADLPFSYFGLVHGRAPAPVFFFTPLVGVGVTMLLATLLIQYE	126		
Qy	128	RRKGVSQSGIMLTPFLWLVALCAJLRSKIMTALKEDAQVDLFRDITFYVYFSLLLIQLV	187		
Db	127	RLOQVSSGVGLIIFWFLFCVWCVAIPVFFRSKILLAKAEGEISDPFRFTFYIFHFALVLSALI	186		

[illegible]



QY 1432 VGORVLCARALLKTKILVLEDEATAAVALDETDLLIQSTIRTFQEDCTVLTIAHRLNTI 1491  
 Db 1428 VGORVLCARALLKTKILVLEDEATAAVALDETDLLIQSTIRTFQEDCTVLTIAHRLNTI 1487  
 QY 1492 MDYTRVILDKGVEIQBYGAPSDLLQORGLFYSAKADAGL 1530  
 Db 1488 MDYTRVILDKGVABEFDSPANLIAARGIFYGNARDAGL 1526

RESULT 9

US-10-295-027-1342  
 ; Sequence 1342, Application US/10295027  
 ; Publication No. US20030232350A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Afar, Daniel  
 ; APPLICANT: Aziz, Nataeha  
 ; APPLICANT: Ginsberg, Wendy M.  
 ; APPLICANT: Gish, Kurt C.  
 ; APPLICANT: Glynn, Richard  
 ; APPLICANT: Hevezi, Peter A.  
 ; APPLICANT: Mack, David H.  
 ; APPLICANT: Murray, Richard  
 ; APPLICANT: Watson, Susan R.  
 ; APPLICANT: Eos Biotechnology, Inc.  
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
 ; FILE OF INVENTION: Methods of Screening for Modulators of Cancer  
 ; FILE REFERENCE: 018501-012500US  
 ; CURRENT APPLICATION NUMBER: US/10/295,027  
 ; CURRENT FILING DATE: 2002-11-13  
 ; PRIOR APPLICATION NUMBER: US 09/663,733  
 ; PRIOR FILING DATE: 2000-09-15  
 ; PRIOR APPLICATION NUMBER: US 60/350,666  
 ; PRIOR FILING DATE: 2001-11-13  
 ; PRIOR APPLICATION NUMBER: US 60/335,394  
 ; PRIOR FILING DATE: 2001-11-15  
 ; PRIOR APPLICATION NUMBER: US 60/332,464  
 ; PRIOR FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: US 60/334,393  
 ; PRIOR FILING DATE: 2001-11-29  
 ; PRIOR APPLICATION NUMBER: US 60/340,376  
 ; PRIOR FILING DATE: 2001-12-14  
 ; PRIOR APPLICATION NUMBER: US 60/347,211  
 ; PRIOR FILING DATE: 2002-01-08  
 ; PRIOR APPLICATION NUMBER: US 60/347,349  
 ; PRIOR FILING DATE: 2002-01-10  
 ; PRIOR APPLICATION NUMBER: US 60/355,250  
 ; PRIOR FILING DATE: 2002-02-08  
 ; PRIOR APPLICATION NUMBER: US 60/356,714  
 ; PRIOR FILING DATE: 2002-02-13  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 1386  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1342  
 ; LENGTH: 1527  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-295-027-1342

Query Match 44.8%; Score 4491.5; DB 15; Length 1527;  
 Best Local Similarity 56.7%; Pred. No. 0;  
 Matches 873; Conservative 260; Mismatches 371; Indels 35; Gaps 9;  
 QY 8 SAGSDPLDMDNVTWNTSPDFTKCFQNTVLVWPCFYLVWACFPFVFLVLSRDRGYIQM 67  
 Db 7 SGEISKFWDSNLSVHTENPDUTPCQNSLLAWPCYLVWALPCYLLVLRHRCGYIIL 66  
 QY 68 TPLNKTALGELLMTVCWADLFYSFWSRSGIFLAPVFLVSPFTLLGITTLATLQIE 127  
 Db 67 SHLSKLMVLGVLLCMVSWADLFYSPHGLVHGRAPVFFVPLVVGVTMLLATLIQYE 126  
 QY 128 RRGVQSSGIMLTFWLVVALVCALATLRSKIMTALKEDAQVDFRDTFFVYPSLLIQIV 187

Db 127 RLQGVQSSGVLIIFWFLCVVCAIVPPRSKILLAKAEISDPFRFTTFFIHFALVLSALI 186  
 QY 188 LSCFSDRSPLPSETIHDNPNCPESASFLSRITFFWITGLIVRGYROPLEGSDLSLNKE 247  
 Db 187 LACFREKPPFFSAKVNDPNYPETAGPLSRIFFFWFTQAIYGYHRLPEEKDLSLKE 246  
 QY 248 DTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKPKESKVDANEEVEALIVKSPQK 307  
 Db 247 DRSQVMVQQLLEWRKQEKQARH-----KASAAPGK---NAGSEDEVLLGARPRP 294  
 QY 308 EWNPSLFKVLKTYTGPYFLMSPFFKAIHDLMMFSPQIILKLIKFNVDTKADQWQCYFYT 367  
 Db 295 R-KPSFLKALLATPGSSFLISACFKLIQDLSFINPQLLSILIRFISNPMPSWGFVA 353  
 QY 368 VLLFVTAQLQTLVHOYFHCIFVSGMRIKTAVIGAVYRKALVITNSARKSSVTGETVNL 427  
 Db 354 GLMFLCSMMQSLILOHYHYIFVTGVKFTGIMGVYRKALVITNSVKRSTVGEIVNL 413  
 QY 428 SVDAQRFMDLATYINNMWSAPLQVILALYLNLNGLSPVLGAVVWVLMVPMVNAVMK 487  
 Db 414 SVDAQRFMDLAPFLNLLSAPLIILATYFLWNLGSPVLGAVFVWLLIPLNGAVAVQM 473  
 QY 488 KTYQVAHMKSKDNRIKLNNEILNGIKVLKYLAWELAFKDKVLAIQBEELKVLKSYLSA 547  
 Db 474 RAFQVKMKLKDSRIKLMSEILNGIKVLKYLAWEPFLKQVREGIRQGELOLLRTAAYLT 533  
 QY 548 VGTFTWCTPPLVALCTFAVYVITDENNLDAOTAFVLSLALFNILRFLNIPMLWISSIV 607  
 Db 534 TTTFTWCMSPPLVLTITLWVYVDPNNVLDKAFVSVLSFNILRPLNMLPQLISNUT 593  
 QY 608 QASVSLKRLRIFLSHEEPDSIERRPVKDGGSITVRNATFTWARSDDPTLNGITFS 667  
 Db 594 QASVSLKRIQQLSQEELDPQSVERTKISPG---YAITHSCTFTWAOQLPFTLSLDT 650  
 QY 668 IPEGALVAVGVQVCGCKSSLISALLAEMDKVEGHVAIGSVAYVPOQAWIQNDSLENTL 727  
 Db 651 VPKGALVAVGVQVCGCKSSLISALLAEMDKVEGHVAIGSVAYVPOQAWIQNDSLENTL 710  
 QY 728 FQCQLEEPYRSVIOACALLPDLPLETPSGDRTEIGEKGWNLGGOKORVSLARAVYSNAD 787  
 Db 711 FGKALNPKRYQOTLEACALLADLEMLPGDQTEIGEKGWNLGGOKORVSLARAVYSNAD 770  
 QY 788 IYLFDDPLSADAHVGKHFENVIGPKMLKNKTRILVTHMSYLPQVDVVIIVMSGGKTS 847  
 Db 771 IFLDPLSADSHVAKHFDHVGPEGVLAKTRVLVTHGIFSLPQTPDFTIIVLADQV 830  
 QY 848 EMGSYQELLARDGAPAEFLRYASTEQODAEENGVTGVSQPKCAKQWENGML----- 901  
 Db 831 EMGPYPALLORNGSPANFLCNYPADPDQGHLEDSWTALEGAEDKEALLIEDTLSNHTLT 890  
 QY 902 ---VTDSSAGLQQLQSSSSSSSGD---ISRHNNSTAB-LQKAEAKKESTWKLMEAD 951  
 Db 891 DNDPVTYVQKQFMQLSALSS-DGEGQGRVPRRHLPSEKQVTEAKADGA--LTQEE 947  
 QY 952 KAQTQGVKLSVWYDMYKAIGLFIPLSIFLFCMNVHSALASNYLSLWTDTPDIPVNGTQSH 1011  
 Db 948 KAAIGTVELSVFWDYAKAVGLCTTILAILLYVQSSAAAIAGANVWLSAWINDAMADSRQN 1007  
 QY 1012 TKVRLSVTGALGISQGIAGVGYSMVAVSISGILASRLHVDLHLSILRSPMSFERTPSGN 1071  
 Db 1008 TSLRLGVYALGILQGLFLVMLAAMAAAGGQAARVLHQAHLHNKIRSPQSFDTTPSGR 1067  
 QY 1072 LVNRPSEKLDVTDSMIPEVIKMFMSLFNVIGACIVILLATPIAAIITPPLGLIYFFVQR 1131  
 Db 1068 ILNCFESKDIYVDEVLAPVILMLNSFFNASTLTVVIMASTPLFTVILPLAVLTLVQR 1127  
 QY 1132 FYVASSRQLKRLSVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYTP 1191  
 Db 1128 FYAATSRQLKRLSVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYTP 1187  
 QY 1192 SIVANRWLWRLCEVGNICVILFAALPAVISHSLSAGLVGLSVLSQVTTVNLMLVRMS 1251  
 Db 1188 YIISNRWLSIGVEFYGNVCLFAALPAVISHSLSAGLVGLSVLSQVTTVNLMLVRMS 1247







Db 1365 PQDPIFGSGLTLMNLDPPFCGRYDEDEIWRTELSHLSAFVSSQPTGLDFQCSEGBDNLVSG 1424  
 Qy 1434 QRLVCLARALKTKILVLDATAVDDLETDLIQSTIRTOFEDCTCTVTIAHRLNTIMD 1493  
 Db 1425 QRLVCLARALKKRSVLVDATAIDAIDLETDLIQSTIRTOFEDCTCTVTIAHRLNTIMD 1484  
 Qy 1494 YTRVIVLDKETOEGVAPSDLLQORGLFYSMAKADAGL 1530  
 Db 1485 YNRVLVLDKGVVAEPDPSVNLIAAGGIFYGMAKADAGL 1521  
 RESULT 11  
 US-10-667-891-1  
 ; Sequence 1, Application US/10667891  
 ; Publication No. US20040171024A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ROTH, CHARLES W.  
 ; APPLICANT: BREY, PAUL T.  
 ; APPLICANT: HOLM, INGE  
 ; APPLICANT: GRAILLES, MARINE  
 ; APPLICANT: RZHETSKY, ANDREY  
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND  
 ; TITLE OF INVENTION: ANOPHELES  
 ; FILE REFERENCE: 03495.0294-00000  
 ; CURRENT APPLICATION NUMBER: US/10/667,891  
 ; CURRENT FILING DATE: 2003-09-23  
 ; PRIOR APPLICATION NUMBER: 60/413,469  
 ; PRIOR FILING DATE: 2002-09-26  
 ; NUMBER OF SEQ ID NOS: 76  
 ; SOFTWARE: Patentin Ver. 3.2  
 ; SEQ ID NO 1  
 ; LENGTH: 1548  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 US-10-667-891-1  
 Query Match 38.5%; Score 3860.5; DB 16; Length 1548;  
 Best Local Similarity 50.0%; Pred. No. 6.2e-318;  
 Matches 784; Conservative 274; Mismatches 438; Indels 71; Gaps 18;  
 Qy 6 FCSADGSDPLDWNVTWNTSNDPDKCFQNTVLVWVPCFYLWA--CFPPFYLYLSHRDG 63  
 Db 12 FCGS----TFWATETWYTNDFDTPCFEQTALVWTPCAFYWAFVDFYLYLKASL-DRN 66  
 Qy 64 YIQMTPLNKTALGFLWVCMADLFYSFWRSGIFLAPVF---LVSPTELLGTTLLA 120  
 Db 67 -IPWKNLVSKALVNLGLVITADLIMALKVG-GDSELPYLDVWGPPIKPAIFLL 124  
 Qy 121 TFLIQLERRKGVSSGIMTLFWLVALVICALAILRSKI-----MTALKEDAQVD 168  
 Db 125 FIFILNKKYGVQTTGCQPIFWFLTLVLSIPRCRTEVRLDAERQKILNSQQPSEQDFSW 184  
 Qy 169 LPRDITFYVYFSLLIQLVLCFSDRSPLFSETIHDNPNCPRESSASFLSRITFWITGLI 228  
 Db 185 EYQFVSFFIFFTFTIMILNCFADGMPQTKYQRGENSEIPLSASFLSRITYQWFDKMA 244  
 Qy 229 VRGYQPLRGSDLSLNKEDTGEQVVPVLVKNWKECAKTRQPKVAVYSSKDPAPKES 288  
 Db 245 LKGYRNPLEKDLNDRLPQDSCSEVNPPIFAHWNQVRKNYKQKARV----EPKAQFNG 300  
 Qy 289 SKVDANEVEALIVKSPQKWNPSLFKVLKTFPGPYFLMSFFPKAIHDLMPFSGPQILKL 348  
 Db 301 NVTFENPHGE---KNGRKKGMASIMPPIYKSGFVGLFGALMKLFTDLTFAQPVLSL 356  
 Qy 349 LKFNVDTKA-PDWQGYFTVLLFVTACLOTVLHQYFHCIFVSGMRKTAIVIGAVYRKA 407  
 Db 357 IISFVEAQDAPEWKGILYAVLLFVLAQAQTFILGQYFHRMFIVGIRINTALINAIYRKA 416  
 Qy 408 LVITNSARKSSTVGETVNLMSVDAOREMDLATYINMWSAPIQVITLALYLLMLNLPSPVL 467  
 Db 417 LRISNTKRESVGEIVNLMAVDAQRFMELTITLYNMWSAPIQIGLALYFLMQQGPSVL 476  
 Qy 468 AGVAVMVLMPVNVAMTKTKYQVAHMKSKONRIKLMNEILNGIKVLKYAWELAFKDK 527

Db 477 AGLAVMIILIPVNGVIASRIKTYQIRQMKYKDERVKLMEVLSGIKVLKYAWERSFEKQ 536  
 Qy 528 VLAIHQEELKVLKSAVLSAVCTFTWCTPPLVALCTFAVYVTDENNILLDACTAFVSLA 587  
 Db 537 VLDIIRDKIEATIARTAYLNAAGTSFLWSCAPPLVSLVTFATYVLIIDENNVLDAKTATKTVS 596  
 Qy 588 LFNILRFPFLNIPMVISSIVQASVSLKRLRPLFLSHSEELPDSIERRRPVKDGCGTNSITVR 647  
 Db 597 LFNILRFPFLTMLPMLITLVLQTVSVNRINKFLNSELDPSNV---LHDSKPKHMSIE 652  
 Qy 648 NATFTWASDPPTLNGITFTSPEGALVAVGVGVGCKSKSLLSALLAEMDKVGHVAIKGS 707  
 Db 653 NGEFSW--GDEITLERNINIEVKKGLVALVGTVGSGKSVVQAFLGEMEKLAGVNVTVGK 710  
 Qy 708 VAYVPOQAWIQNDSURENIFCGQLEEPYRSVIOACALLPDLLEILPSGDRTEIGKGVN 767  
 Db 711 LAYVPOQAWIQNATVRDNLFGQTYDRKRYNKVIOACALRADIDILSAGDLTEIGKGIN 770  
 Qy 768 LSGGQKQVSLARAVYSNADIYLFDDPLSAVDAAHVGHKIFENVIGPKMLKNKTRILVTH 827  
 Db 771 LSGGQKQVSLARAVYSADLYLLDDPLSAVDAAHVGHKIFERVIGPKGLARKSVLVT 830  
 Qy 828 SMSYLPQVDVIVMSGGKISEMGSYQELLARDGAPAEFLRTYASTEQEODAEEN----- 881  
 Db 831 GVTFELPQVDSIVYIKWGEISESGTDPQLVKNKGAFADFIQHLQEGNEEBEELNQIKRQI 890  
 Qy 882 ---GVTGVSQPKQKQWENGMLVTDSE-----AGKQLORQLSSSSSYSGDISR 926  
 Db 891 SSTADVPBELLGTVEKAIKLARTESLSDSISVTSADSLMGGGGSLRRRTKQDSS----- 943  
 Qy 927 HHNSTAEQLKAEAKKEETWKLMEADKAQTGOVKLSVYDYMKAIGLIFISFLIFMCMNH 986  
 Db 944 -HDSVASAASLKKQKQEVGKLIETEKSTQGGVEFAVYKHYIKSVGIFLSVATLVNLFVQ 1002  
 Qy 987 VSALASNYWLSLWTDPIVNGTQEHKTVRLSVYGALGISQGIYAVGYSNAVSGIGILASR 1046  
 Db 1003 AFQIGSNLWLTQWANDQNVANDTGLRDMYLVGYGAFGQVLSKYLGLALAIIGGLHCSM 1062  
 Qy 1047 CLHVDLLSHLSIRSPMSFFERTPSGNLVRFSKELDTVDSMIPEVIMKPMGSLFNVIAGCI 1106  
 Db 1063 NVFNKLLNTGLKWPMLFDTPLGRILSKYSKVDVTVDSVLPAITVQLLNTCFGLATTIV 1122  
 Qy 1107 VILLATPIAAIIIPPLGLIYFFVQRYFVASSRQLARLESVSRSPYVSHFNELTLLGVSVIR 1166  
 Db 1123 VISLSTPIFLAVIVPIAFIYFAQRYFVATSRQLARLESVSRSPYVSHFSETVGTASTIR 1182  
 Qy 1167 APEQERTIHOSDLKVDENQKAYTPSIIVANRWLAVLECVGNCIVLPAALFAVIRSHSLS 1226  
 Db 1183 AYNVGDRFIESDAKVDKNQVCKYPSVIANRWLAIEMVGNLIILFASLFAVLGGQT-N 1241  
 Qy 1227 AGLVGLSVSYSLQVTTYLNWLVRMSSEMETNIVAVERLKEYSETKEAPWQI-QETAPPS 1285  
 Db 1242 PELVGLSVSYALQVQTYLNWLVRMSSDIETNLVSVRIKEYGETKQEAPEWLEQDKNKKP 1301  
 Qy 1286 SWPQVRFEFRNYCLIRYREDLDFVLRHINVTINGCKEYVIGVTRGAGKSSLTGLFRINE 1345  
 Db 1302 NWPQGRVEFQNFQRYREGDLVLRGVSNFIQCGEKVIGVTRGAGKSSLTGLFRITE 1361  
 Qy 1346 SAEGBIITDGINIAKIGHDLRFKTIIPQDPVLPSGSLRMLNLDPPSQVSDSEVTSLEL 1405  
 Db 1362 AAGGRISIDGVDIASMGLMLRSLRTIIPQDPVLPSGSLRMLNLDPPFETKTDDEIKWALE 1421  
 Qy 1406 AHLKDFVSALPDKDHECAEGENLSVGORQLVCLARALLRKTILVLDEATAAVDLETD 1465  
 Db 1422 SHLSFVLSLAAGLNEHIEABGENLSVGORQLVCLARALLRKTILVLDEATAAVDLETD 1481  
 Qy 1466 DLIQSTIRTOFEDCTCTVTIAHRLNTIMDYTRVIVLDKGEIOEYGAAPSLLQO-RGLFYSM 1524  
 Db 1482 DLIQKTIETFEKCTVLTIAHRLNTILSDKVIIVLDKQIIEFASFTTELLDNPKSAFYSM 1541  
 Qy 1525 AKDAGLV 1531  
 |||||

```

Db      1542 AKDANLV 1548

RESULT 12
US-10-667-891-4
; Sequence 4, Application US/10667891
; Publication No. US20040171024A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, CHARLES W.
; APPLICANT: BREY, PAUL T.
; APPLICANT: HOLM, INGE
; APPLICANT: GRAILLES, MARINE
; APPLICANT: RZHETSKY, ANDREY
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
; TITLE OF INVENTION: ANOPHELES
; FILE REFERENCE: 03495.0294-00000
; CURRENT APPLICATION NUMBER: US/10/667,891
; CURRENT FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: 60/413,469
; PRIOR FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patentin Ver. 3.2
; SEQ ID NO 4
; LENGTH: 1505
; TYPE: PRT
; ORGANISM: Anopheles gambiae
US-10-667-891-4

Query Match      36.0%; Score 3603.5; DB 16; Length 1505;
Best Local Similarity 48.0%; Pred. No. 5.9e-296;
Matches 750; Conservative 280; Mismatches 444; Indels 87; Gaps 22;

QY      1 MALRFGCSADGSDPLDWNWNTWNTGNPDTKCFQNTVLVWVPCFVLMACFPFYLX-LSR 59
Db      1 WTFDEFQCG-----PFWDFGEFVDVNDPNLTFCFQVILQWPCLEFV-PSIYDFXITE 55
QY      60 HDRQVIQMTPLNKTALGFLFLMWVCADL-FYSFWSRSGIFLAPVFLVSPTLIGITTL 118
Db      56 SKYRIDPWNWNLRSMLVFLMLCMCIDLGMVGYQDBEGLYDQVILTAFAVNALAYIDL 115
QY      119 LATELIQELRRKGVOSSGMLTFWLVALVCAALILRSKIM-----TALKEDAQVDL--F 170
Db      116 LV--LFPWRKIGVATSGTMFMFLRPFGLIQLRTEVMENDKRPNAIGSGDVTDFWEY 173
QY      171 RDTFYVYFSLLIQLIOLSCFSDRSPLFSETHDNPCESSASPLSRITFWMITGLIVR 230
Db      174 QVVSYLQVSLICMLVLELPDKEPSFYYPKAAKPNELRSFFSKLLFLHFDAPAWK 233
QY      231 GYRQPLEGSDLSWLNKEDTSQVVPVLVKNWKECAKTRKQPVKVYVSSKDPAPQKSSK 290
Db      234 GFRNPLTNDMYDINPODSARELVPFPDKYWKISVEKGRKQOM-----ASDRKAGK 284
QY      291 VDANEEVEALIVKSPQKEMNPSLFKVLKTF----GPFYFLMSFFFKAIHDLMMFSGQIL 346
Db      285 PDID-----YKPHSFNSGSLVLTMTAVGGPFWFAGMLQLAISGL-QFASPYLM 332
QY      347 KLLIKFVNDTKAPDQGGFYFVTLVFTACLOTLVLHQYFHFICFVSGMRKIKTAVIGAVRK 406
Db      333 QELMAVI-AFDGPLWKGPGLLTFGLFGASLLGLFNGQVLFYFVFLSGFRIRGLISALYRK 391
QY      407 ALVITNSARKSTVEIENLMSVDAQRPMDLATYINMIWSAPLOVIALYALYLLWNLGFSV 466
Db      392 ALRISSAAKDTTVGEIENLMAVDAQRFELTSLYHLHLSALLIIGLCVFLLYDILGPV 451
QY      467 LAGVAVMVLVMPVNAVMAKTKTYQVAHMKSKDNRIKLMNBILNGIKVLKLYAWELAFKD 526
Db      452 FAGLGWMLTVPVSGVAAKLKTHQVQAKLKDERVKKMKNEILGGIKVLKLYAWEPFQD 511
QY      527 KVLAIROBELKVLKKSAYLSAVGTFTWVCTPFLVALCTFAVYVITDENNILDQAQAFVSL 586
Db      512 SILNVRDEEVGILKMAVYAGIETFTIAPFLVLSFAVYVILIDENNVLDPQTAQAFVSL 571
QY      587 ALFNILRFPNLPWISSIVQASVSLKRLIRIFLSHEELPDSIERPVPKGGGTNSITV 646

Db      572 ALFNIMRFPGLGMPFPMVVTFSQAWYSIKRIDKFLNSAELDPNNVTHNK-----SDEALTI 626
QY      647 RNATFTWARSDDPTLNGITFSPISPEGALVAVGVCGCGKSSLLSALLAEWDKVEGHVAIKG 706
Db      627 KDGTFESWG-DETPTLKNLNLKQKLSAIVGTGTGKSSLLSALLGEMEKISGHVNTDG 685
QY      707 SVAYVPOQAWIQNDSLRNILLFGCQLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKG 766
Db      686 SIAYVPOQAWIQNATLRDNLILFGKAFDQKDYNDVIECCALRPDLLEMLPGDSTEIGEGI 745
QY      767 NLSGGOKORVSLARAVYSNADIYLPDDPLSAVDHVGHIIPENVIGPKGMLKNKTRILVT 826
Db      746 NLSGGOKORVALARAVYADAAYVYLPDDPLSAVDHVGKHIPEKIVGPSGMLVKGSRLLVT 805
QY      827 HSMYSYLPQVDVIIVMSSGKISGMSGYQELLARDGAFAEFLRTYASTEQEOADAEENGVTGV 886
Db      806 HGISFLPVENILVLKDEISESGTYQELIDQKGAFAEFLSQHL---QELDDEDEISLI 862
QY      887 SGPGKEAKQOMENGLVTDTSAGKQIQORQLSSSSSYSG-----DISRH-----HNST 931
Db      863 Q-----ETLNDGVV-----NNVIQALSVRSNRSNGSDGSTRKKPISRQVSKQSVHKT 911
QY      932 AELQKAEAKKEETWKLMEADKAQTGOVKLSYVWDMKAIGLFISSLIFLFCMKNHVSALA 991
Db      912 VTTVPGRA-----NLIGVEESATGAVTWLVYKXIYIOSIGPKFGFVFLTAIINOSSGIF 965
QY      992 SNYLSLWTDPIVNGTQEHKTKVRLSVYALGISQGIAGVFGYSMAVSTGGTILASRCLHVD 1051
Db      966 SNLWTDWSEDPDAATDPSVRDKYLVGYVAGLGAQSIALFVAALLISIGCLKAAKESHNK 1025
QY      1052 LLHSILSPMSFPFRTSGNHLVNRFSKELTDVDSMIPEVIMKFMGSLFNVTGACIVILLA 1111
Db      1026 LLESCLRMPSFFDTPLGRIINRFSKDVVDVNDVNLVPTIRAWLFLFNFGVFIIGTS 1085
QY      1112 TPIAAIIIPGLGLTYFFQVFPYVASSRQLKRLSVSRSPVYSHFNETHLIGVSVIRAPPEEQ 1171
Db      1086 TPIFLAVVPLVWLYYVQRFYIDTSQKRLSVTRSPYSHFGESIGGOSTIRAYCQQ 1145
QY      1172 ERFHQSOLDKVDENQKAYPSIVANRWLAVLECVGNCIVLFAALFAVIHRHLSAGLVG 1231
Db      1146 DRFTQESERRVDVNLQVSYPTIVANRWLAVLELIGSCVILFAALFAILARDTIGQATVG 1205
QY      1232 LSVSYSLQVTTYLNWLMVMSMETNIVAVRELKEYSETEKAPWQIOETAPSSWQVQ 1291
Db      1206 VSIYALQISHLVSLVMTSEVETNIVAVLEEYITVLPREAEW---OKGTVDKAWPVBG 1263
QY      1292 RVEFRNYCLRYREDLDFVLRRHINVTINGGKVGIVGTGACKSSLTGLFLFRINESARGEI 1351
Db      1264 KVEFKDYQIRYREGDLIVIRGISLVNRGGEKIGIVGTGACKSSLTGLFLFRIVEAAGQI 1323
QY      1352 IIDGINITAKILHDLRPKTIIPQDPVLFSGSLRMNLDPPFSQYSDDEBWTSLSLAHUKDF 1411
Db      1324 IIDGLDISKMLGLQLGRLLTIIPQDPVLFSGTLRANVDPFKSYSDDOVKWKAELSHLTKTF 1383
QY      1412 VSALPDKLDHECAEGENLSVGORQVLCARALLRKTILVLEATAVLETDLDTLQST 1471
Db      1384 VKGLSAGLDHEIAENGENLSVGORQLICLARAVLRKTVLILDEATAVLETDLDTLQST 1443
QY      1472 IRTOPEDECTVLTIAHRLNTIMDYTRVIVLDKGBIOEYGAPESDLLQOR-GLFYSMADKAGL 1530
Db      1444 IRTEFADCTLTIAHRLNTILSDRVLDKGLVAECDSQNLNANRESIFYGMKNAGI 1503
QY      1531 V 1531
Db      1504 V 1504

```

RESULT 13  
 US-10-667-891-3  
 ; Sequence 3, Application US/10667891  
 ; Publication No. US20040171024A1  
 ; GENERAL INFORMATION:



```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1542
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-363-112-4

```

Query Match		35.9%	Score 3600;	DB 15;	Length 1542;
Best Local Similarity		47.6%;	Pred. No. 1.2e-295;		
Matches 747;	Conservative 287;	Mismatches 462;	Indels 72;	Gaps 20;	

  

QY	3	LRGFCSDGGDPLDMKNVTNTS-----NPDFTKCFONTVLWVPCFLWACFPFYFL--	55
Db	2	LEKECN-----STFWNSSFDSPEADPLCFEQTIVLWITPLGFLMLAPWQLHV	51
QY	56	YLSRHDRGYIQMTPLNKTKTALGELLIVCWADLFYSFWSRSRGIFLAPVLPSPHTLLGI	115
Db	52	YKSRTKRSSTTKLYLAK-QVFVGELL-LAAIAELAVLTEDSGOATPVAVRYTNPSLY-L	108
QY	116	TTLATATELIOLERRKGVOSSGIMLT-FWLVALVCALAILRSKINTALKEDAQVDLFRDIT	174
Db	109	GTMLLVULIQQSRQWCQKNSWFSLSEWIIISLCGTFQFOITRLTQG-DNSNLAYSCL	167
QY	175	FYYVFSLLLQLVLSCFSDRSPFSETHPNPCESSASFLSRITFWWTIGTLVRGVQ	234
Db	168	FFISYGFOILLIFSASFEN-----NESSNNPSSIASFLSITTSWDVSILKYKGR	219
QY	235	PLGSDLMNLINKEDTSQVPVLVGNKKCEAKTRK-----QPVKVYSSKD	281
Db	220	PLTLEDVWEDEEMKTTLVSKFETHMKRELQARRALORRQBKSSQONSQCARLPGLNKN	279
QY	282	PAOPKESSKDVAEEVEALLVKGPKQEWPNF--LFKVLYKTFGPYFLMSFEFAIHDLMM	339
Db	280	QSOSQDALVL---EDVEKKKKSGTKDVDPKSWLMKALFKTFYMWLKSFLKLVDNIFT	336
QY	340	FSPQIILIKLIKFNVDTKADPWOGYFVVLLFVTACLQTLVLHQYPHICFVSGMRKTAIV	399
Db	337	FVSPQLKLLISPASRDYTLWIGYLCAILLFTAALITSQCLQCYFQCFLKGKVRTAI	396
QY	400	IGAVYRKALVITNSARKSVTGEIVNLSVDAQRFMDLATYINMISAPLOVITALYLW	459
Db	397	MASYKKAULTSLNLAKEYITGTVTNLSVDAQLMQMDVTNFHMLMSSVQIVLSITFFLW	456
QY	460	LNLGPSVLGAVAVMVLMPVNAMKTKTYQVAHMKSNDNRILKLANEILNGIKVLKLYA	519
Db	457	RELGPSVLAGVVMVLVIPINALISTKSKIQVKNMKNDKRLKIMNEILLSGILKYFA	516
QY	520	WELAFDKVLAIRQEBLVLKSAIYLSAVCTFTWCTPFVALCTFAYVIIDENNILDA	579
Db	517	WPFSFDQVQNLRKSELNLLAFSQOCVVFVQTFPVLVSVTVSVVYLVDSNNILDA	576
QY	580	QTAFVSLALFNILRPFLNTPILPMVISIVOASVSLKELRIPLFSHEELEPDISIERRPVKDG	639
Db	577	QKRAFTITLNLRRFLPMLPMMISSMLQASVSTERLEYLGGLDDLDTSAIRH----DCN	632
QY	640	GTNSITVRNATFWARSDDPTLNGITFSPGHALVAVQVCGCKGSLSALLAEADMKYE	699
Db	633	FKMAQFSEASFWEHDSBATVRDNLDTMAGQLVAVIPGVSGSKSLISAMLGENVNH	692
QY	700	GHAIVKGSVAYYPQQAWIQNDSLRENILFGCOLLEEPPYRSVIOACALLPLELTPSGDRT	759
Db	693	GHITIKTATTAYPQSQWINGTIKONILGFTEFENEKRYQQVLEACALLPLEMLPGCDLA	752
QY	760	EITEKGVNLSGGOKQRVSLARAVYSNADYILFDDPLSAVDADVHGKHIFENVIGPKOMLKN	819
Db	753	EITEKGINLSGGOKQRISLARATYQNLDTIYLLDDPLSAVDADVHGKHIFNKVLGPNLLKG	812
QY	820	KTRLIVTHSMSVLPQVDVIIVMSGGKISBMGYSOEILLARDGAFAEFLRTY-ASTEOEQDA	878
Db	813	KTRLIVTHSMHFLPOVDIEIVLNGTIVZKGSYSALLAKKGEFAKMLKTLFLRHTGPEESA	872
QY	879	-----ENGVTGVSGPQKEAKQMENGMLVTDTSAGKIQRLSGSSSYSGDISRHNST	931

## RESULT 15

```

RES001.13
US-10-252-155-4
; Sequence 4, Application US/10252155
; Publication No. US20040068096A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS IN ORGANIC ANION TRANSPORT
; TITLE OF INVENTION: MULTI-DRUG RESISTANT PROTEINS
; FILE REFERENCE: D0152 NP
; CURRENT APPLICATION NUMBER: US/10/252,155
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US 60/324,172
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/333,700
; PRIOR FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 793
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1545
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-252-155-4

```

Query Match 35.9%; Score 3600; DB 15;  
Best Local Similarity 47.6%; Pred. No. 1.2e-295;  
Matches 747; Conservative 287; Mismatches 462; Indels 72;



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:19 ; Search time 139.718 Seconds  
(without alignments)  
5389.603 Million cell updates/sec

Title: US-10-665-283-8

Perfect score: 10016

Sequence: 1 MALRGFCAGSDPLMDWNV.....RSVAVAKAKPFSIPDLSL 1947

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7860	78.5	1531	2	AAW57486 Human MRP
2	7860	78.5	1531	2	AAW74471 Human mul
3	7860	78.5	1531	2	AAW99894 Human mul
4	7860	78.5	1531	3	AAV55799 Human mul
5	7860	78.5	1531	3	AAV78873 Multidrug
6	7860	78.5	1531	5	ABG61810 Prostate
7	7860	78.5	1531	6	ABM35012 Cancer ba
8	7860	78.5	1531	6	ABD20865 MRP1 base
9	7860	78.5	1531	7	ADB87954 Human UGT
10	7860	78.5	1531	7	ADB96937 Human MDR
11	7860	78.5	1531	7	ADB92128 Human MDR
12	7860	78.5	1531	7	ADD44764 Human Pro
13	7860	78.5	1531	8	ADN97111 MRP1 prot
14	7856	78.4	1530	7	ADN95929 Human BEC
15	7849	78.4	1531	2	AAV54928 Multidrug
16	7849	78.4	1531	2	AAV93153 Multi-dru
17	7849	78.4	1531	2	AAW57485 Human mul
18	7849	78.4	1531	2	AAW74470 Human mul
19	7849	78.4	1531	2	AAW99893 Human mul
20	7849	78.4	1531	3	AAV55798 Human mul
21	7849	78.4	1531	3	AAV78872 Human mul
22	7849	78.4	1531	3	AAV93152 Multi-dru
23	7810	76.0	1489	2	AAV96952 Multi-dru
24	7587.5	75.8	1482	8	ADN03902 Antipsori
25	7587.5	75.8	1482	8	ADP23422 PRO polyp

26	7572	75.6	1481	5	ABP52108 Homo sapi
27	7560.5	75.5	1482	2	AAV96953 Multi-dru
28	7500.5	74.9	1472	2	AAV96954 Multi-dru
29	7230.5	72.2	1530	7	ADP56611 Bovine MR
30	7194	71.8	1417	2	AAV96955 Multi-dru
31	7037	70.3	1388	7	ADJ69912 Human hea
32	7002.5	69.9	1528	2	AAW57487 Murine mu
33	7002.5	69.9	1528	2	AAW74472 Mouse mul
34	7002.5	69.9	1528	2	AAW99895 Mouse mul
35	7002.5	69.9	1528	3	AAV55800 Murine mu
36	7002.5	69.9	1528	3	AAV78874 Murine mu
37	6916	69.0	1372	2	AAV98211 Multi-dru
38	6911	69.0	1345	2	AAV96956 Multi-dru
39	4491.5	44.8	1527	5	AAU91309 Human pro
40	4491.5	44.8	1527	5	ABP52110 Homo sapi
41	4491.5	44.8	1527	6	ABR58639 Human can
42	4491.5	44.8	1527	7	ADP54387 Human pro
43	4491.5	44.8	1527	7	ADN40024 Cancer/an
44	4491.5	44.8	1527	8	ADP24088 PRO polyp
45	4487.5	44.8	1527	2	AAV43543 A human M

## ALIGNMENTS

RESULT 1  
AAW57486  
ID AAW57486 standard; protein; 1531 AA.  
XX  
AC AAW57486;  
XX  
DT 14-AUG-1998 (first entry)  
XX  
DE Human MRP variant ltpgpa (Lei/Pgpa).  
XX  
KW Multidrug resistance-associated protein; MRP; tumour; human; variant;  
KW multidrug resistance; MDR; leishmania P-glycoprotein; ltpgpa; Lei/Pgpa.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 685 /label= L685S  
FT /note= "wild-type Leu is replaced by Ser"  
FT Misc-difference 1282 /label= R1282A  
FT /note= "wild-type Arg is replaced by Ala"  
XX

US5766880-A.  
16-JUN-1998.  
05-JUN-1995; 95US-00463092.  
27-OCT-1992; 92US-00966923.  
08-MAR-1993; 93US-00029340.  
26-OCT-1993; 93US-00141893.  
20-MAR-1995; 95US-00407207.  
(TOOH ) UNIV QUEENS KINGSTON.  
Cole SP, Deeley RG;  
WPI; 1998-361687/31.  
N-PSDB; AAV31498.

DNA encoding protein associated with multi-drug resistance - useful for as probe for identifying multi-drug resistant tumour cells.  
Claim 1; Col 67-78; 82pp; English.  
This represents a variant of the human multidrug resistance-associated protein (MRP). This natural variant is a leishmania P-glycoprotein related

CC molecule lcpGpA (Lei/Pgpa). The human and murine MRP nucleic acid  
 CC molecules can be used as probes for identifying multidrug resistant  
 CC tumour cells by hybridisation to mRNA from tumour cells. The antisense  
 CC nucleic acid can be used to reverse multidrug resistance (MDR). A  
 CC recombinant expression vector containing the MRP nucleic acid molecules  
 CC operatively linked to at least one regulatory sequence can be used to  
 CC transform a host cell to produce a recombinant MDR-associated protein  
 XX  
 SQ Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 2; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGCSDGSDPLMDWNVNTNTSNPDFTKCFQNTVLWVPCFYLWACFPFYLILSRH 60  
 DB 1 MALRGCSDGSDPLMDWNVNTNTSNPDFTKCFQNTVLWVPCFYLWACFPFYLILSRH 60

QY 61 DRGYIQTMTPLNKTALGFLWVWADLFYFWSRSGIFLAPVPLVSPILGITTLLA 120  
 DB 61 DRGYIQTMTPLNKTALGFLWVWADLFYFWSRSGIFLAPVPLVSPILGITTLLA 120

QY 121 TFLIOLRRKGVSIGIMLTFWLVALCALAIRSKIMTALKEDAQVDFRDIIFYVYFS 180  
 DB 121 TFLIOLRRKGVSIGIMLTFWLVALCALAIRSKIMTALKEDAQVDFRDIIFYVYFS 180

QY 181 LLLIOLVLSGSDRPLSETIHDNCPPESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LLLIOLVLSGSDRPLSETIHDNCPPESSASFLSRITFWITGLIVRGYRQPLEGSD 240

QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL 300  
 DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL 300

QY 301 IVKSPQKWNPSLFKVLKTPGYPLMGFFPKAIHDLWVFGPQILKLIKFVNDTKAPD 360  
 DB 301 IVKSPQKWNPSLFKVLKTPGYPLMGFFPKAIHDLWVFGPQILKLIKFVNDTKAPD 360

QY 361 WQGYFTVLLFVTACQLVLHVOYPHICFVSGMRKTAIVIGAVYKALVITNSARKSTV 420  
 DB 361 WQGYFTVLLFVTACQLVLHVOYPHICFVSGMRKTAIVIGAVYKALVITNSARKSTV 420

QY 421 GEIVNLMSVDAQRFMDLTYINIMWSAPLQVILALYLLWMLGSPVLGAVVWVLPVN 480  
 DB 421 GEIVNLMSVDAQRFMDLTYINIMWSAPLQVILALYLLWMLGSPVLGAVVWVLPVN 480

QY 481 AVMAKTKTYQVAHMKSDKNRIKLMNEILNGIKVLKYAMELAFKDKVLAIROBELKVLK 540  
 DB 481 AVMAKTKTYQVAHMKSDKNRIKLMNEILNGIKVLKYAMELAFKDKVLAIROBELKVLK 540

QY 541 KSAYLSAVGTFTWCTPFLVALCTPAVVTVDENNILDQAOTAFVSLALFNILRPLNLP 600  
 DB 541 KSAYLSAVGTFTWCTPFLVALCTPAVVTVDENNILDQAOTAFVSLALFNILRPLNLP 600

QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGTSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGTSITVRNATFTWARSDDPT 660

QY 661 LINGITFSPGALVAVVQVCGCKSSLLSALLAEMDKVEGHVAKGSVAVYVQQAQWTD 720  
 DB 661 LINGITFSPGALVAVVQVCGCKSSLLSALLAEMDKVEGHVAKGSVAVYVQQAQWTD 720

QY 721 SLRENILFGCOLERPYRSVIOACALLPDLRILPSGDRTEIGEGVNLSSGQKORVSLAR 780  
 DB 721 SLRENILFGCOLERPYRSVIOACALLPDLRILPSGDRTEIGEGVNLSSGQKORVSLAR 780

QY 781 AVYNADILYFDDPLSAVDHVGXHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 DB 781 AVYNADILYFDDPLSAVDHVGXHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840

QY 841 MSGGKISEMGVSQELLARDGAFELRTYASTEOBDAENGVTGVSQPGKEAKOMENG 900  
 DB 841 MSGGKISEMGVSQELLARDGAFELRTYASTEOBDAENGVTGVSQPGKEAKOMENG 900

QY 901 LVTDAGKQLOLQSSSSSYSGDISRHNSHTAELOKAEAKKEBTWKLMEADKAQTQGVKL 960  
 DB 901 LVTDAGKQLOLQSSSSSYSGDISRHNSHTAELOKAEAKKEBTWKLMEADKAQTQGVKL 960

QY 961 SVYDYMKAIGLFTSIFLIFCMCNHVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 DB 961 SVYDYMKAIGLFTSIFLIFCMCNHVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020

QY 1021 ALGISQGIHAFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTSGNLVNRFSKEL 1080  
 DB 1021 ALGISQGIHAFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTSGNLVNRFSKEL 1080

QY 1081 DTVDMSIPEVIMKFMGSLFNIVIGACIVILLATPIAAIIIPBLGLIYFFVQRPVYASSROL 1140  
 DB 1081 DTVDMSIPEVIMKFMGSLFNIVIGACIVILLATPIAAIIIPBLGLIYFFVQRPVYASSROL 1140

QY 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYSIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYSIVANRWLA 1200

QY 1201 VRLECVCNIVLFAALFAVISHRSLSAGLVSVSYSLQVTTYLNLVMSSEMSEMETNIVA 1260  
 DB 1201 VRLECVCNIVLFAALFAVISHRSLSAGLVSVSYSLQVTTYLNLVMSSEMSEMETNIVA 1260

QY 1261 VERLKEYSETEKEAPWQIQTETAPPSSWPQVGRVFRNCLRYREDLDFVLRHINVTINGG 1320  
 DB 1261 VERLKEYSETEKEAPWQIQTETAPPSSWPQVGRVFRNCLRYREDLDFVLRHINVTINGG 1320

QY 1321 EKVGIVGRTGAGKSLTGLFRINESAGEEIIIDGINIAKIGLHDLRKFITIIIPDDPVLF 1380  
 DB 1321 EKVGIVGRTGAGKSLTGLFRINESAGEEIIIDGINIAKIGLHDLRKFITIIIPDDPVLF 1380

QY 1381 SGSLRMNLDPFSQVSDSEVWTSLELAHLKDFVSALPKLDHECAEGGNSLVGQRLVCL 1440  
 DB 1381 SGSLRMNLDPFSQVSDSEVWTSLELAHLKDFVSALPKLDHECAEGGNSLVGQRLVCL 1440

QY 1441 ARALLRKTILVLEDEATAVDDLETDLLIQTSTRTQFEDCTVLTIAHRLNTIMDYTRVIL 1500  
 DB 1441 ARALLRKTILVLEDEATAVDDLETDLLIQTSTRTQFEDCTVLTIAHRLNTIMDYTRVIL 1500

QY 1501 DKGIEQYCAPSDLLQQRGLFYSMAXDAGLV 1531  
 DB 1501 DKGIEQYCAPSDLLQQRGLFYSMAXDAGLV 1531

RESULT 2  
 AAW74471  
 ID AAW74471 standard; protein, 1531 AA.  
 XX  
 AC AAW74471;  
 XX  
 DT 18-MAY-1999 (first entry)  
 XX  
 DE Human multidrug resistance-associated protein variant.  
 XX  
 KW Multidrug resistance-associated protein; MDR; human; diagnosis;  
 KW MDR tumour cell identification; cancer therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FN US5882875-A.  
 XX  
 PD 16-MAR-1999.  
 XX  
 PF 05-JUN-1995; 95US-00462109.  
 XX  
 PR 27-OCT-1992; 92US-00966923.  
 PR 08-MAR-1993; 93US-00029340.  
 PR 26-OCT-1993; 93US-00141893.  
 PR 20-MAR-1995; 95US-00407207.  
 XX  
 PA (TOOH ) UNIV QUEENS KINGSTON.



XX Cole SPC, Deeley RG;  
 XX WPI: 1999-214061/18.  
 XX N-PSDB; AAX21977.  
 PT Identifying a multidrug resistant tumor cell by contacting the cell with  
 PT an antibody/antigen-binding fragment - which binds to an expressed  
 PT protein encoded by multidrug resistance-associated protein (MRP) nucleic  
 PT acid.  
 XX  
 XX Claim 3; Col 69-80; 80pp; English.  
 XX  
 CC This sequence is the human multidrug resistance-associated (MDR) protein.  
 CC The invention relates to a method for identifying a multidrug resistant  
 CC (MDR) tumour cell. Compositions and methods utilising the MDR proteins  
 CC can be used to treat patients with tumours displaying multidrug  
 CC resistance, particularly those displaying resistance to anthracyclines,  
 CC epipodophylotoxins, vinca alkaloids, and hydrophobic drugs. The methods  
 CC for inhibiting/killing a MDR tumour cell can be useful for treating  
 CC breast cancer, leukaemias, fibrosarcomas, cervical cancer, gliomas,  
 CC thymomas, neuroblastomas and lung cancer. The MDR DNA sequences when  
 CC labeled are useful as molecular probes for diagnosing multidrug  
 CC resistance of a tumour (using cells from a tumour biopsy) and for  
 CC designing ribozymes which are capable of cleaving a single-stranded  
 CC nucleic acid encoding a protein having MRP activity. Recombinant  
 CC expression vectors containing human MDR coding sequences can be  
 CC transfected into a drug sensitive cell line to produce a protein in the  
 CC cell which confers MDR, protecting non-resistant non-tumour cells from  
 CC the effects of chemotherapeutics has major clinical importance. Cells  
 CC transformed with the MDR coding sequences are useful for testing  
 CC potential therapeutic agents for their effectiveness against MDR cells  
 CC and for identifying chemosensitisers of a therapeutic agent  
 XX  
 XX Sequence 1531 AA;  
 SQ  
 Query Match 78.5%; Score 7860; DB 2; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAIRGFCSDGSDPLDWDNVTWNTSNPDTKCFQNTVWVPCFYLWACFPFYLYLSRH 60  
 DB 1 MAIRGFCSDGSDPLDWDNVTWNTSNPDTKCFQNTVWVPCFYLWACFPFYLYLSRH 60  
 QY 61 DRGYQMTPLNKTALGFLMIVCWADLFYFWSRSGIFLAPVFLSPDLLGTTLLA 120  
 DB 61 DRGYQMTPLNKTALGFLMIVCWADLFYFWSRSGIFLAPVFLSPDLLGTTLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVICALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TFLIQLERRKGVQSSGIMLTFWLVALVICALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 QY 181 LLLIQLVLSCFDRSPLSETTHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LLLIQLVLSCFDRSPLSETTHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPQKESKVDANEVEAL 300  
 DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPQKESKVDANEVEAL 300  
 QY 301 IVKSPQKWNPSLFKVLKXTFGPYFLMSFFFKAIHDLMMFSGPQILKLIKFVNDTKAPD 360  
 DB 301 IVKSPQKWNPSLFKVLKXTFGPYFLMSFFFKAIHDLMMFSGPQILKLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFTACLOTLLVHOYFHCIFVSGMRKTAIVGAYRKALVITNSARKSTV 420  
 DB 361 WQGYFYTVLLFTACLOTLLVHOYFHCIFVSGMRKTAIVGAYRKALVITNSARKSTV 420  
 QY 421 GRIVNLSMDAORFMDLATVINNIWSAPLOQVILALYLNLGSPSLAGVAVVLMVFN 480  
 DB 421 GRIVNLSMDAORFMDLATVINNIWSAPLOQVILALYLNLGSPSLAGVAVVLMVFN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLAENILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540

481 AVMAKTKTYQVAHMKSKDNRIKLAENILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFVSLALFNILRFPNLNLP 600  
 DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTDENNILDAQTAFVSLALFNILRFPNLNLP 600  
 QY 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKDGSGTNSITVRNATFTWARSDDPT 660  
 QY 661 LINGITFISPEGALVAVVQVCGGKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720  
 DB 661 LINGITFISPEGALVAVVQVCGGKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720  
 QY 721 SLRENILFGCOLLEBPYYSVIOACALLPDLELTPSGDRTEIGEGKGNVLSGGQKQVSLAR 780  
 DB 721 SLRENILFGCOLLEBPYYSVIOACALLPDLELTPSGDRTEIGEGKGNVLSGGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIIV 840  
 DB 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIIV 840  
 QY 841 MSGGKISEMSYQELLARDGAPAEPLRTYASTEQDQDAEENGVTGVSFGPKAKOMENGM 900  
 DB 841 MSGGKISEMSYQELLARDGAPAEPLRTYASTEQDQDAEENGVTGVSFGPKAKOMENGM 900  
 QY 901 LVTSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAEAKKEETWKLMEADKAOQGVKL 960  
 DB 901 LVTSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAEAKKEETWKLMEADKAOQGVKL 960  
 QY 961 SVYDYMKAIGLIFISFLIFLMCNHVSALASNYWLSLWTDDDPIVNGTQEHKTVRLSVYG 1020  
 DB 961 SVYDYMKAIGLIFISFLIFLMCNHVSALASNYWLSLWTDDDPIVNGTQEHKTVRLSVYG 1020  
 QY 1021 ALGISQGIATVFGYSMAVSIIGGILASRCCLHVDLLHSILRSPMSFFERTPSGNLVNRFPSKEL 1080  
 DB 1021 ALGISQGIATVFGYSMAVSIIGGILASRCCLHVDLLHSILRSPMSFFERTPSGNLVNRFPSKEL 1080  
 QY 1081 DTVDMSIPEVIMKFMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 DB 1081 DTVDMSIPEVIMKFMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 QY 1201 VRLECVGNICVILFAALPAVISHRSLSAGLVLSVSYSLQVTTYLNLVVRMSSEMETNIYA 1260  
 DB 1201 VRLECVGNICVILFAALPAVISHRSLSAGLVLSVSYSLQVTTYLNLVVRMSSEMETNIYA 1260  
 QY 1261 VERLKEYSETKEAPWQIQETAPSSWPQVGRVPRNYCLRYREDLDVFLRHINVTINGG 1320  
 DB 1261 VERLKEYSETKEAPWQIQETAPSSWPQVGRVPRNYCLRYREDLDVFLRHINVTINGG 1320  
 QY 1321 EKVGVIGRTGAGKSSLTGLFRINESARGEIIIDGINIAKIGHDLRPFKTIIPQDPVLF 1380  
 DB 1321 EKVGVIGRTGAGKSSLTGLFRINESARGEIIIDGINIAKIGHDLRPFKTIIPQDPVLF 1380  
 QY 1381 SGSLRMNLDPPSQVSDERBWTSLAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 DB 1381 SGSLRMNLDPPSQVSDERBWTSLAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 QY 1441 ARALRKTILVLDATAVLDLTDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
 DB 1441 ARALRKTILVLDATAVLDLTDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
 QY 1501 DKGEIQEYCAPSDLLQORGLFYSMADAGLV 1531  
 DB 1501 DKGEIQEYCAPSDLLQORGLFYSMADAGLV 1531

AAW99894  
 ID AAW99894 standard; protein; 1531 AA.  
 XX  
 AC AAW99894;  
 AC  
 DT 10-JUN-1999 (first entry)  
 DT  
 XX  
 DE Human multidrug resistance-associated protein natural variant.  
 XX  
 KW Human; multidrug resistance-associated protein; MRP; cytotoxic drug;  
 KW cancer; chemotherapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5891724-A.  
 XX  
 PD 06-APR-1999.  
 XX  
 PF 05-JUN-1995; 95US-00460907.  
 XX  
 PR 27-OCT-1992; 92US-00966923.  
 PR  
 PR 08-MAR-1993; 93US-00029340.  
 PR  
 PR 26-OCT-1993; 93US-00141893.  
 PR  
 PR 20-MAR-1995; 95US-00407207.  
 XX  
 XX (TOOH ) UNIV QUEENS KINGSTON.  
 PA  
 XX  
 PI Cole SPC, Deeley RG;  
 XX  
 DR WPI, 1999-253868/21.  
 DR  
 DR N-PSDB; AAX19818.  
 XX  
 XX Protecting mammalian cells against cytotoxic drugs.  
 PT  
 XX  
 PS Claim 2; Col 79-86; 82pp; English.  
 XX  
 CC The present sequence represents a human multidrug resistance-associated  
 CC protein (MRP). The present invention also describes a method for  
 CC protecting a mammalian cell against the cytotoxicity of anthracyclines,  
 CC epipodophyllotoxins and Vinca alkaloids (A) by introducing into it a  
 CC nucleic acid (I) that hybridizes under stringent conditions to a nucleic  
 CC acid (II) that encodes an MRP protein (III). Introduction of (I) protects  
 CC cells against cytotoxic effects of (A), particularly to protect normal  
 CC cells against (A) being used for treatment of cancers. Cells transformed  
 CC with (I) can be used to screen for agents that affect multidrug  
 CC resistance or are directly toxic to multidrug resistant cells, i.e.  
 CC potential therapeutics for multidrug-resistant cancers. Confering  
 CC resistance to normal cells should allow an increase in the dose of (A)  
 CC that can be administered safely  
 XX  
 SQ Sequence 1531 AA;  
 Query Match 78.5%; Score 7860; DB 2; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGFCADGSDPLMDNWNWNTSNPDTKCFQNTLVWVPCFYLWACPFYFLYLSRH 60  
 DB 1 MALRGFCADGSDPLMDNWNWNTSNPDTKCFQNTLVWVPCFYLWACPFYFLYLSRH 60  
 QY 61 DRGYIQMTPLNKTALGFLWIVCWADLFYSFWSRSGIFLAPVFLVSPFLGITTLLA 120  
 DB 61 DRGYIQMTPLNKTALGFLWIVCWADLFYSFWSRSGIFLAPVFLVSPFLGITTLLA 120  
 QY 121 TFLIQLEKRGVQSSGIMLTFLWALVCAALILRSKIMTALKEDAQVDLPDIFFYVYFS 180  
 DB 121 TFLIQLEKRGVQSSGIMLTFLWALVCAALILRSKIMTALKEDAQVDLPDIFFYVYFS 180  
 QY 181 LLLIQVLSCDSRSLPSETIHPNCPSSASFLSRITFWWTGTIIVRGYRPLRGSD 240  
 DB 181 LLLIQVLSCDSRSLPSETIHPNCPSSASFLSRITFWWTGTIIVRGYRPLRGSD 240  
 QY 241 LWSLNKEDTSQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPKSSKVDANEEVEAL 300

DB 241 LWSLNKEDTSQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPKSSKVDANEEVEAL 300  
 QY 301 IVKSPQKEWNPSLFKVLYKTFPGYFLMGSFFPKAIHDLMMFSGPQILKLLIKFVNNTKAPD 360  
 DB 301 IVKSPQKEWNPSLFKVLYKTFPGYFLMGSFFPKAIHDLMMFSGPQILKLLIKFVNNTKAPD 360  
 QY 361 WQGYFYTVLLFVTTACLTQTLVHLQYFPHICFVSGMRITKTAIVIGAVYRKALVITNSARKSSTV 420  
 DB 361 WQGYFYTVLLFVTTACLTQTLVHLQYFPHICFVSGMRITKTAIVIGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLMSVDAQRFMDLATYINMIWSADLOVTLALYLIMNLGSPVLGAVVIMLVPNV 480  
 DB 421 GEIVNLMSVDAQRFMDLATYINMIWSADLOVTLALYLIMNLGSPVLGAVVIMLVPNV 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPPLNLP 600  
 DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLALFNILRFPPLNLP 600  
 QY 601 MVISSIVOASVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVOASVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 QY 661 LINGITFSIPEGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
 DB 661 LINGITFSIPEGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
 QY 721 SLRENILFGCOLLEPYRYSVIOACALLPDLLEILPSGDRTEIGEKGVLNLSGQKQKRVSLAR 780  
 DB 721 SLRENILFGCOLLEPYRYSVIOACALLPDLLEILPSGDRTEIGEKGVLNLSGQKQKRVSLAR 780  
 QY 781 AVYSNADIYLPDDPLSAVDAHVGHKIPENVIGPKGMLKNKTRILIVTHSMSVLPQVDVILV 840  
 DB 781 AVYSNADIYLPDDPLSAVDAHVGHKIPENVIGPKGMLKNKTRILIVTHSMSVLPQVDVILV 840  
 QY 841 MSGGKISEMGYSQBELLDARDAFAEFLRTYASTEQDAENGVTGVSQPGKEAKQEMNGM 900  
 DB 841 MSGGKISEMGYSQBELLDARDAFAEFLRTYASTEQDAENGVTGVSQPGKEAKQEMNGM 900  
 QY 901 LVTDSAGKQIQORQSSSSSSSGDLSRHNSHSTAELOKAEKEETWKLMEADKQAGOVKL 960  
 DB 901 LVTDSAGKQIQORQSSSSSSSGDLSRHNSHSTAELOKAEKEETWKLMEADKQAGOVKL 960  
 QY 961 SVYWDYMKAIGLFISFLSIFLFCNHNVSALASNYLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 DB 961 SVYWDYMKAIGLFISFLSIFLFCNHNVSALASNYLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 QY 1021 ALGISQGIAGVGYSMAYSIGILLASRCLHVDLHLSILRSPMSFFERTPSGNLVRNRFKEL 1080  
 DB 1021 ALGISQGIAGVGYSMAYSIGILLASRCLHVDLHLSILRSPMSFFERTPSGNLVRNRFKEL 1080  
 QY 1081 DTVDSDMIPVYKFMFMSLNFVIGACIVILLATPAAIIIPPLGLIYFFVORFYVASSRQL 1140  
 DB 1081 DTVDSDMIPVYKFMFMSLNFVIGACIVILLATPAAIIIPPLGLIYFFVORFYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYGHFNETLLGVSVIRAFEPQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYGHFNETLLGVSVIRAFEPQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 QY 1201 VRLECVCNCILVFAALFAVLSRHSLSAGLVLSVSYSLQVTTVYLNWLRMSSEMETNIVA 1260  
 DB 1201 VRLECVCNCILVFAALFAVLSRHSLSAGLVLSVSYSLQVTTVYLNWLRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKAPMQIQETAPPPSWPQVGRVFEFNYCLIRYREDLDFVLRHINVTINGG 1320  
 DB 1261 VERLKEYSETEKAPMQIQETAPPPSWPQVGRVFEFNYCLIRYREDLDFVLRHINVTINGG 1320  
 QY 1321 EKVGIVGRGAGKSSITGLGFRINESABGEIIGDGINAKIGIHLDRFKITTIPODPVLV 1380

Db 1321 EKVGIVGRTGAGKSLTGLFRINSAEGBEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLSRNLNDPFFQYSDDEEVTWTSLELAHLKDFVSALPDKLDHCAEGGNSLVGQRLVCL 1440  
 Db 1381 SGLSRNLNDPFFQYSDDEEVTWTSLELAHLKDFVSALPDKLDHCAEGGNSLVGQRLVCL 1440  
 QY 1441 ABALLRKTKILVLDREATAVLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ABALLRKTKILVLDREATAVLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIOEYCAPSDLLQOORGLFYSMAKDAGLV 1531  
 Db 1501 DKGEIOEYCAPSDLLQOORGLFYSMAKDAGLV 1531

RESULT 4

AAV55799  
 ID AAV55799 standard; protein; 1531 AA.  
 XX  
 AC AAV55799;  
 XX  
 DT 28-FEB-2000 (first entry)  
 XX  
 DE Human multidrug resistance-associated protein (MRP) variant.  
 XX  
 KW Chemosensitizer; multidrug resistance-associated protein; MRP; human;  
 KW therapeutic agent; P-glycoprotein-mediated multidrug resistance; lung;  
 KW cancer; variant.  
 XX  
 OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Misc-difference 685 /label= L685S  
 FT /note= "wild-type Leu is replaced by Ser"  
 FT Misc-difference 1282 /label= R1282A  
 FT /note= "wild-type Arg is replaced by Ala"

XX US6001563-A.  
 XX PD 14-DEC-1999.  
 XX PF 05-JUN-1995; 95US-00463179.  
 XX PR 27-OCT-1992; 92US-00966923.  
 XX PR 08-MAR-1993; 93US-00029340.  
 XX PR 26-OCT-1993; 93US-00141893.  
 XX PR 20-MAR-1995; 95US-00407207.  
 XX PA (TOOH ) UNIV QUEEN'S KINGSTON.  
 XX PI Cole SP, Deeley RG;  
 XX WPI; 2000-061877/05.  
 XX DR N-PSDB; AA239556.  
 XX PT Identification of chemosensitizers useful for treating cancer, using  
 XX nucleic acids encoding multidrug resistance-associated protein.  
 XX PS Claim 3; Col 69-80; 77pp; English.  
 XX CC The invention provides a method for identifying a substance which is a  
 XX chemosensitizer that comprises, contacting a cell transfected with  
 XX nucleic acid encoding multidrug resistance-associated protein (MRP) with  
 XX a therapeutic agent in vitro. The method is useful for identifying  
 XX chemosensitizers which may then be used to treat cancer (especially lung  
 XX cancer). The method allows the identification of chemosensitizers which  
 XX do not reverse P-glycoprotein-mediated multidrug resistance. The present  
 XX sequence represents a human MRP variant  
 XX SQ Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 3; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGFCSDGSDPLDNDNNTWNTSNPDFTKCFQNTVLVWVPCFVLMACFPFYLYLSRH 60  
 Db 1 MALRGFCSDGSDPLDNDNNTWNTSNPDFTKCFQNTVLVWVPCFVLMACFPFYLYLSRH 60  
 QY 61 DRGYQMTPLNKTKTALGFLWLWVCWADLFYSFWSRSRGIFLAPVFLVSPDLLGITLLA 120  
 Db 61 DRGYQMTPLNKTKTALGFLWLWVCWADLFYSFWSRSRGIFLAPVFLVSPDLLGITLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFVYFS 180  
 Db 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFVYFS 180  
 QY 181 LLLIQLVLSGSDRSPLESETHDNPSPSSASFLSRITFWITGLIVRGVROPLEGSD 240  
 Db 181 LLLIQLVLSGSDRSPLESETHDNPSPSSASFLSRITFWITGLIVRGVROPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKOPKVYVYSSKOPAPQKSSKVDANEVEAL 300  
 Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKOPKVYVYSSKOPAPQKSSKVDANEVEAL 300  
 QY 301 IVKSPQKEWNPFLFKVLYKTGPFYFLMSFFFKAIHDLMMFSGPQILKLIKFNVDTKAPD 360  
 Db 301 IVKSPQKEWNPFLFKVLYKTGPFYFLMSFFFKAIHDLMMFSGPQILKLIKFNVDTKAPD 360  
 QY 361 WQGYFYTLLFVTAQLOTLVHLQYFHCIVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTLLFVTAQLOTLVHLQYFHCIVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINMWSAPLOVTLALYLWNLGSPSLAGVAVMLMVPVN 480  
 Db 421 GEIVNLSVDAQRFMDLATYINMWSAPLOVTLALYLWNLGSPSLAGVAVMLMVPVN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYLAWELAFKDKVLAIQBELKVLK 540  
 Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYLAWELAFKDKVLAIQBELKVLK 540  
 QY 541 KSAYLSAVGTTFWCTPFLVALCTFAVVTVDENNILDAQTAFAVLSALFNILRFPNLIIP 600  
 Db 541 KSAYLSAVGTTFWCTPFLVALCTFAVVTVDENNILDAQTAFAVLSALFNILRFPNLIIP 600  
 QY 601 MVISSIVOASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660  
 Db 601 MVISSIVOASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660  
 QY 661 LNGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPQAWIQND 720  
 Db 661 LNGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPQAWIQND 720  
 QY 721 SLRENILFGCOLEEPYRSVIOACALLPDLETLPSGDRTEIGEGVNLGGOKQVSLAR 780  
 Db 721 SLRENILFGCOLEEPYRSVIOACALLPDLETLPSGDRTEIGEGVNLGGOKQVSLAR 780  
 QY 781 AVYSNADIIYLFDDPLSADVAHVGHKIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 Db 781 AVYSNADIIYLFDDPLSADVAHVGHKIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFAPLRTYASTEQEQDAEENGVTGVSGPGKEAKQENGM 900  
 Db 841 MSGGKISEMGSYQELLARDGAFAPLRTYASTEQEQDAEENGVTGVSGPGKEAKQENGM 900  
 QY 901 LVTSAGLQLOLORQLSSSSSYSGDISRHNSIAELOKABAKKEETWKLMEADKAQGVKL 960  
 Db 901 LVTSAGLQLOLORQLSSSSSYSGDISRHNSIAELOKABAKKEETWKLMEADKAQGVKL 960  
 QY 961 SVYDYMKAIGLFTSIFLIFMCMNVHSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 Db 961 SVYDYMKAIGLFTSIFLIFMCMNVHSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 QY 1021 ALGISQGIATVFGYSMAVSGIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSL 1080

1021 ALGISQIAVFGYSMAVSGIGLILASCLHVDLHLSILRSPMFETPTSGNLVNRFSKEL 1080  
 1081 DTVDMSIPEVIMKFMGSLFNIGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSRQL 1140  
 1081 DTVDMSIPEVIMKFMGSLFNIGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSRQL 1140  
 1141 KRLESVSRSPVSHNETILLGVSIVIRAFEEQERFHQSOLKVDENKAYPSIVANRWLA 1200  
 1141 KRLESVSRSPVSHNETILLGVSIVIRAFEEQERFHQSOLKVDENKAYPSIVANRWLA 1200  
 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGSLVSYSLOVTTYLNLVMSSEMETNIVA 1260  
 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGSLVSYSLOVTTYLNLVMSSEMETNIVA 1260  
 1261 VRLKEYSTEKEAPQIQTETAPPSWPQVGRVFRNRYCLRYREDLDFVLRHINTVINGG 1320  
 1261 VRLKEYSTEKEAPQIQTETAPPSWPQVGRVFRNRYCLRYREDLDFVLRHINTVINGG 1320  
 1321 EKVGIVRTGAGKSSLTGLFRINSAEGEIIIDGINIAKIGLHDLRFKTIIPDDPVL 1380  
 1321 EKVGIVRTGAGKSSLTGLFRINSAEGEIIIDGINIAKIGLHDLRFKTIIPDDPVL 1380  
 1381 SGSLRNLDPPFQYSDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 1381 SGSLRNLDPPFQYSDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 1441 ARALARLKTILVLDATAAVLETDLQSTIRTQPEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 1441 ARALARLKTILVLDATAAVLETDLQSTIRTQPEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 1501 DKGEIQEYGAPSDLLQQRGLFYSMADAGLV 1531  
 1501 DKGEIQEYGAPSDLLQQRGLFYSMADAGLV 1531

## RESULT 5

AAY78873  
 ID AAY78873 standard; protein; 1531 AA.

AC AAY78873;

DT 19-MAY-2000 (first entry)

DE Multidrug resistance protein (MRP) natural variant amino acid sequence.

DE Multidrug resistance protein; MRP; human; anthracycline; vinca alkaloid;  
 KW epipodophyllotoxin; cancer; leukaemia.

XX Homo sapiens.

OS US6025473-A.

PN 15-FEB-2000.

PD 05-JUN-1995; 95US-00461384.

PF 27-OCT-1992; 92US-00966923.

PR 08-MAR-1993; 93US-00029340.

PR 26-OCT-1993; 93US-00141893.

PR 20-MAR-1995; 95US-00407207.

XX (TOOH ) UNIV QUEENS KINGSTON.

XX Cole SPC, Deeley RG;

XX WPI; 2000-181838/16.

XX N-PSDB; AAZ90193.

XX Isolated protein conferring multidrug resistance, to at least two drugs  
 PT selected from anthracyclines, epipodophyllotoxins and vinca alkaloids, on  
 PT a drug sensitive mammalian cell.

XX

PS Claim 10; Col 79-88; 78pp; English.

XX This sequence represents a human multidrug resistance protein (MRP)  
 CC natural variant amino acid sequence. The human MRP confers multidrug  
 CC resistance, including resistance to at least two drugs selected from  
 CC anthracyclines, epipodophyllotoxins and vinca alkaloids, on a drug  
 CC sensitive mammalian cell, when the protein is expressed in the cell. The  
 CC multidrug resistance is not substantially reversed by chemosensitizers  
 CC which reverse P-glycoprotein-mediated multidrug resistance. The MRP  
 CC protein sequence can be used to generate antibodies against MRP. The MRP  
 CC protein and nucleotide sequences can be used in compositions which are  
 CC used to treat patients with tumours displaying multidrug resistance. The  
 CC compositions and methods of the invention can be used particularly to  
 CC treat breast cancer, leukaemias, fibrosarcomas, cervical cancer, gliomas,  
 CC thymomas, neuroblastomas, and lung cancer. Antibodies directed against  
 CC MRP can be used to inhibit the multidrug resistance of a multidrug  
 CC resistant cell

SQ Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 3; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLDNDNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFYFLYSRH 60  
 DB 1 MALRGFCSADGSDPLDNDNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFYFLYSRH 60  
 QY 61 DRGVIQMTPLNKTALGFLIIVCWADLFYSFERSRGIFLAPVELVSPFLIGITTLA 120  
 DB 61 DRGVIQMTPLNKTALGFLIIVCWADLFYSFERSRGIFLAPVELVSPFLIGITTLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILSKIMTALKEDAQVDLFRDITFYVVS 180  
 DB 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILSKIMTALKEDAQVDLFRDITFYVVS 180  
 QY 181 LLLIQLVLSCFSDRSPFLSETIHDNCPSSASFSLRITFWMITGLIVRGYRPLSGSD 240  
 DB 181 LLLIQLVLSCFSDRSPFLSETIHDNCPSSASFSLRITFWMITGLIVRGYRPLSGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANEEVEAL 300  
 DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANEEVEAL 300  
 QY 301 IVKSPQKEWNPFLFKVLYKTFGYPFLMSFPFKAIHDLMPFSGPOILKLLIKFVNDTKAPD 360  
 DB 301 IVKSPQKEWNPFLFKVLYKTFGYPFLMSFPFKAIHDLMPFSGPOILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTACLTQLVLHGVFHCIFVSGMRITKTAIVGAVYRKALVITNSARKSSTV 420  
 DB 361 WQGYFYTVLLFVTACLTQLVLHGVFHCIFVSGMRITKTAIVGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINNIWSAPLQVILALYLLNLNGLSPVLGAVMVLMPVN 480  
 DB 421 GEIVNLSVDAQRFMDLATYINNIWSAPLQVILALYLLNLNGLSPVLGAVMVLMPVN 480  
 QY 481 AVNAMKTKTYQVAHMSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROEELKVLK 540  
 DB 481 AVNAMKTKTYQVAHMSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROEELKVLK 540  
 QY 541 KSAVLSAVGTFTWVCTPFLVALCTFAVYVITDENNILDATQAFVSLALFNILFPLNLP 600  
 DB 541 KSAVLSAVGTFTWVCTPFLVALCTFAVYVITDENNILDATQAFVSLALFNILFPLNLP 600  
 QY 601 MWISSIVQASVSLKRLRIFLSHEELPDSIERPPVKDGGTNSITVRNATFTWARSDDPT 660  
 DB 601 MWISSIVQASVSLKRLRIFLSHEELPDSIERPPVKDGGTNSITVRNATFTWARSDDPT 660  
 QY 661 LNCITTSIPGALVAVVGQVCGKSKLLSALLAEMDKVEGHVAIKGSVAVVPOQMTQND 720  
 DB 661 LNCITTSIPGALVAVVGQVCGKSKLLSALLAEMDKVEGHVAIKGSVAVVPOQMTQND 720  
 QY 721 SURENIFGQCUEEPYRYSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGQKQVSLAR 780

Db 721 SLRENILFCQLEPYRYSVIOACALLPDLEILPSGDRTEIGEGVNLSSGGQKQVSLAR 780  
 Qy 781 AVYSNADIYLFDDPLSAVDAHVGKHFENVIKPKMLKNKTRILVTHSMSYLPQVDVIV 840  
 Db 781 AVYSNADIYLFDDPLSAVDAHVGKHFENVIKPKMLKNKTRILVTHSMSYLPQVDVIV 840  
 Qy 841 MSGKISEMGYSQELLARDGAFELRTYASTEQEQDAEENGVTGVSFGKKEAKQWENG 900  
 Db 841 MSGKISEMGYSQELLARDGAFELRTYASTEQEQDAEENGVTGVSFGKKEAKQWENG 900  
 Qy 901 LVTDSAGKOLQRLSSSSYSYSGDISRHNSHSTAELOKAEAKKEETWKLMEADYAOQTQV 960  
 Db 901 LVTDSAGKOLQRLSSSSYSYSGDISRHNSHSTAELOKAEAKKEETWKLMEADYAOQTQV 960  
 Qy 961 SVYDYMKAIGLFIISFLSIFLFCMCHVLSALSNYLSLWTDPIVNGTOEHTKVLRSYVG 1020  
 Db 961 SVYDYMKAIGLFIISFLSIFLFCMCHVLSALSNYLSLWTDPIVNGTOEHTKVLRSYVG 1020  
 Qy 1021 ALGISQIAVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMFFERTPSGNLVNRFKSEL 1080  
 Db 1021 ALGISQIAVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMFFERTPSGNLVNRFKSEL 1080  
 Qy 1081 DTVDMSIPVIMKMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 Db 1081 DTVDMSIPVIMKMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 Qy 1141 KRLESVSRSPVYSHFNETLLGVSIRAPPEQERFIHQSDLVKVDENOKAYPSIVANRWLA 1200  
 Db 1141 KRLESVSRSPVYSHFNETLLGVSIRAPPEQERFIHQSDLVKVDENOKAYPSIVANRWLA 1200  
 Qy 1201 VRLECVGNCILVFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNWLRMSSEMETNIVA 1260  
 Db 1201 VRLECVGNCILVFAALFAVISRHSLSAGLVLSVSYSLQVTTYLNWLRMSSEMETNIVA 1260  
 Qy 1261 VERLKEYSTEAPWQIETAPPSPQVGRVFRNYCLRYREDLDFVLRHINTYINGG 1320  
 Db 1261 VERLKEYSTEAPWQIETAPPSPQVGRVFRNYCLRYREDLDFVLRHINTYINGG 1320  
 Qy 1321 EVKGVIGRTGAGKSSLTGLFRINSAEAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 Db 1321 EVKGVIGRTGAGKSSLTGLFRINSAEAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 Qy 1381 SCSLRNLDPPFSQYSDVEEVTSLAHLKDFVSALPKDLHCEAGGENLSYQRLVCL 1440  
 Db 1381 SCSLRNLDPPFSQYSDVEEVTSLAHLKDFVSALPKDLHCEAGGENLSYQRLVCL 1440  
 Qy 1441 ABALLRKTKILVLDATAVDETDLLIOSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ABALLRKTKILVLDATAVDETDLLIOSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Qy 1501 DKGEIQEYCAPSDDLQOQRLGFYSMAKDAGLV 1531  
 Db 1501 DKGEIQEYCAPSDDLQOQRLGFYSMAKDAGLV 1531

RESULT 6

ID ABG61810 standard; protein; 1531 AA.  
 XX AC ABG61810;  
 XX DT 15-AUG-2002 (first entry)  
 XX DE Prostate cancer-associated protein #11.  
 XX KW Prostate cancer; prostate tumour tissue; human; mammal; cytotatic.  
 XX OS Mammalia.  
 XX PN WO20020268-A2.  
 XX PD 18-APR-2002.

XX 12-OCT-2001; 2001WO-US032045.  
 XX 13-OCT-2000; 2000US-00687576.  
 PR 08-DEC-2000; 2000US-00733288.  
 PR 08-DEC-2000; 2000US-00733742.  
 PR 24-JAN-2001; 2001US-0263957P.  
 PR 16-MAR-2001; 2001US-0276791P.  
 PR 16-MAR-2001; 2001US-0276888P.  
 PR 06-APR-2001; 2001US-0281922P.  
 PR 24-APR-2001; 2001US-0286214P.  
 PR 30-APR-2001; 2001US-00847046.  
 PR 04-MAY-2001; 2001US-0288589P.  
 XX (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;  
 WPI; 2002-471335/50.  
 XX N-PSDB; ABK92125.  
 DR Detecting a prostate cancer-associated transcript in a cell in a patient,  
 PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,  
 PT by determining if prostate cancer-associated genes are expressed in a  
 PT prostate tissue.  
 XX Claim 27; Page 309; 436pp; English.  
 XX The present invention relates to methods of detecting a prostate cancer-  
 CC associated transcript in a cell from a patient. The method comprises  
 CC contacting a biological sample from the patient with prostate cancer-  
 CC associated polynucleotides (designated PC genes) that selectively  
 CC hybridise to a sequence that is at least 80% identical to them. The  
 CC prostate cancer-associated polynucleotide sequences are differentially  
 CC expressed in prostate tumour tissue or in prostate cancer and are derived  
 CC from the tissues of various organisms such as humans or other mammals  
 CC (e.g. mice, sheep and dogs). The methods of the invention are useful for  
 CC diagnosing and treating prostate cancer in mammals. The prostate cancer-  
 CC associated genes are useful for diagnosing or treating prostate cancer,  
 CC as well as for identifying modulators of prostate cancer or agents that  
 CC inhibit prostate cancer. The nucleic acid sequences are particularly  
 CC useful in gene therapy, as a vaccine or in antisense applications.  
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins.  
 XX Sequence 1531 AA;  
 SQ  
 Query Match 78.5%; Score 7860; DB 5; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MALRGFCSADGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFVLWACFPFYLYLSRH 60  
 Db 1 MALRGFCSADGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFVLWACFPFYLYLSRH 60  
 Qy 61 DRGYQMTPLNKTALGFLWIVCWADLFYSFWSRSRGIFLAPVFLVSPITLLGTTLLA 120  
 Db 61 DRGYQMTPLNKTALGFLWIVCWADLFYSFWSRSRGIFLAPVFLVSPITLLGTTLLA 120  
 Qy 121 TFLQLERRKGVQSSGIMLTFWLVLCALAILRSKINTALKEDAQVDLFRDITFVYVFS 180  
 Db 121 TFLQLERRKGVQSSGIMLTFWLVLCALAILRSKINTALKEDAQVDLFRDITFVYVFS 180  
 Qy 181 LLLQLVLSCFSDRSPLFSETHDPNCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
 Db 181 LLLQLVLSCFSDRSPLFSETHDPNCPSSASFLSRITFWITGLIVRGVQPLEGSD 240  
 Qy 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKQPKVYVSSKDPAPQKSSKVDANEVEAL 300  
 Db 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKQPKVYVSSKDPAPQKSSKVDANEVEAL 300  
 Qy 301 IVKSPQKEWNSLKFVLYKTGPFVLSMFFPKAHDLMFSGPQTLKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKEWNSLKFVLYKTGPFVLSMFFPKAHDLMFSGPQTLKLLIKFVNDTKAPD 360

QY 361 WQGYFTVLLFVTAQLQTLVHOFHICFVSGMRIKTAIVIGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFTVLLFVTAQLQTLVHOFHICFVSGMRIKTAIVIGAVYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLMSVDAQRFMDLATVINNMWSAPLQVILALYLLNLGSPSLAGVAVVNLMPVN 480  
 Db 421 GEIVNLMSVDAQRFMDLATVINNMWSAPLQVILALYLLNLGSPSLAGVAVVNLMPVN 480  
 QY 481 AVAMKTKYQVAHMSKONRIKLNNEILNGIKVLYAWELAFKDKVLAIRQBELKVLK 540  
 Db 481 AVAMKTKYQVAHMSKONRIKLNNEILNGIKVLYAWELAFKDKVLAIRQBELKVLK 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVITDENNILDQAQFVSLALNIRFPLNILP 600  
 Db 541 KSAYLSAVGTFTWCTPFLVALCTFAVYVITDENNILDQAQFVSLALNIRFPLNILP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERPVKGGGTNSITVRNATFTWASDPPT 660  
 Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERPVKGGGTNSITVRNATFTWASDPPT 660  
 QY 661 LMGITFSIPEGALVAVVGQVCGKSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720  
 Db 661 LMGITFSIPEGALVAVVGQVCGKSLLSALLAEMDKVEGHVAIKGSVAYVPOQAWIQND 720  
 QY 721 SLRENILFCQLEEPYRSVIOACALLPDLILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
 Db 721 SLRENILFCQLEEPYRSVIOACALLPDLILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
 QY 781 AYSNADIYLPDPLISAVDAHVGKIHIFENVIOPKGMKKNKTRILVTHSMSYLPQVDVITV 840  
 Db 781 AYSNADIYLPDPLISAVDAHVGKIHIFENVIOPKGMKKNKTRILVTHSMSYLPQVDVITV 840  
 QY 841 MSGGKISEMGYSQELLARDGAEFLRTYASTEQDAEENGVTGSGPGKEAKOMENG 900  
 Db 841 MSGGKISEMGYSQELLARDGAEFLRTYASTEQDAEENGVTGSGPGKEAKOMENG 900  
 QY 901 LVTDSAGKQQLQSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQGTQVKL 960  
 Db 901 LVTDSAGKQQLQSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQGTQVKL 960  
 QY 961 SVYDYMKAIGLIFSLFSLFPMCHNVSNALSNYLSLWTDPIVNGTQEHKTVRLSYG 1020  
 Db 961 SVYDYMKAIGLIFSLFSLFPMCHNVSNALSNYLSLWTDPIVNGTQEHKTVRLSYG 1020  
 QY 1021 ALGISQGIAGVGYMAVSIIGTILASRCLHVLHLSILRSPMSFFERTPSGNLVRFSKEL 1080  
 Db 1021 ALGISQGIAGVGYMAVSIIGTILASRCLHVLHLSILRSPMSFFERTPSGNLVRFSKEL 1080  
 QY 1081 DTVDNMIPEVIMFMGSLFNVTGACIVILLATPTAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 Db 1081 DTVDNMIPEVIMFMGSLFNVTGACIVILLATPTAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 QY 1141 KRLSVSRSPVSHNETLLGVSVIRAEQERFTHQSDLVKDNOKAYYSIVANRWLA 1200  
 Db 1141 KRLSVSRSPVSHNETLLGVSVIRAEQERFTHQSDLVKDNOKAYYSIVANRWLA 1200  
 QY 1201 VRLCEVGNICVLFAALFAVIRSRHLSAGLVLSYSYLSQVTTYLNWLRMSSEMETNIVA 1260  
 Db 1201 VRLCEVGNICVLFAALFAVIRSRHLSAGLVLSYSYLSQVTTYLNWLRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETKEAPWQIQTAPSPSPWQGRVFRNRYCLYRDLDPVLRHINVTINGG 1320  
 Db 1261 VERLKEYSETKEAPWQIQTAPSPSPWQGRVFRNRYCLYRDLDPVLRHINVTINGG 1320  
 QY 1321 EKVGVGTGTAGKSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 Db 1321 EKVGVGTGTAGKSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGSLRMNLDPPSQSVDEEVTSLAHLKDFVSALPKDLHECAEGGENLSVGQRLVCL 1440  
 Db 1381 SGSLRMNLDPPSQSVDEEVTSLAHLKDFVSALPKDLHECAEGGENLSVGQRLVCL 1440

QY 1441 ARALLRKTKILVLDATAVDDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTKILVLDATAVDDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQEYGAPELDLQQRGLFYSMADAGLV 1531  
 Db 1501 DKGEIQEYGAPELDLQQRGLFYSMADAGLV 1531

RESULT 7  
 ABM35012  
 ID ABM35012 standard; protein; 1531 AA.  
 XX  
 AC ABM35012;  
 XX  
 DT 08-OCT-2003 (first entry)  
 XX  
 DE Cancer based on CYP3A5 related protein SEQ ID NO:678.  
 XX  
 KW Cancer; CYP3A5; irinotecan; pharmaceutical; malignant glioma;  
 KW cytochrome p450; subfamily IIIA; nifedipine oxidase; polypeptide 5;  
 KW cytostatic.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2003013534-A2.  
 XX  
 PD 20-FEB-2003.  
 XX  
 PF 23-JUL-2002; 2002WO-EP008219.  
 XX  
 PR 23-JUL-2001; 2001EP-00117608.  
 PR 24-MAY-2002; 2002EP-00011710.  
 XX  
 PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
 XX  
 PI Heinrich G, Kerb R;  
 XX  
 DR WPI; 2003-268144/26.  
 XX  
 PT New use of irinotecan for preparation of compositions for treating cancer  
 PT in subject having genome with variant allele comprising cytochrome p450,  
 PT subfamily IIIA, polypeptide 5 polynucleotide, termed CYP3A5.  
 XX  
 PS Disclosure; SEQ ID NO 678; 86pp; English.  
 XX  
 CC The present invention describes the use of irinotecan (I) or its  
 CC derivative for the preparation of a pharmaceutical composition for  
 CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic  
 CC cancer, or malignant glioma in a subject having a genome with a variant  
 CC allele which comprises a cytochrome p450, subfamily IIIA (nifedipine  
 CC oxidase), polypeptide 5 (CYP3A5) polynucleotide (II). (I) and (II) have  
 CC cytostatic activity. The therapeutic applications of (I) is improved,  
 CC since it is possible to individually treat a subject with an appropriate  
 CC dosage and/or an appropriate derivative of (I). Therefore, undesirable,  
 CC harmful or toxic effects are efficiently avoided. Unnecessary and  
 CC potentially harmful treatment of those subjects who do not respond to the  
 CC treatment with substances (nonresponders), as well as the development of  
 CC drug resistances due to suboptimal drug dosing can be avoided. ACFe2200  
 CC to ACFe2751 and ABM34912 to ABM35013 represent sequences used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 6; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCADSGDPLMDNVTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFYLILSRH 60  
 Db 1 MALRGFCADSGDPLMDNVTWNTSNPDFTKCFONTVLVWVPCFYLWACFPFYLILSRH 60  
 QY 61 DRGYIQTPLNKTKTALGFLNLIVCWADLFSYFWSRSGIFLAPFLVSLTLLGITTLLA 120

Db 61 DRGYQMTPLNKTALGELLIVCWADLFYFWERSRGIFLAPVLPSPPTLLGTTLLA 120  
 QY 121 TELIOLERRKGVOSSGIMLTFWLVALVLCALAILRSKIMTALKEDAQVDLFRDITFVYVPS 180  
 Db 121 TELIOLERRKGVOSSGIMLTFWLVALVLCALAILRSKIMTALKEDAQVDLFRDITFVYVPS 180  
 QY 181 LLLIOLVLSFCSDRSPFSETIHDNPNCPDESSASFLSRITFWIITGLIVRGYRQPLEGSD 240  
 Db 181 LLLIOLVLSFCSDRSPFSETIHDNPNCPDESSASFLSRITFWIITGLIVRGYRQPLEGSD 240  
 QY 241 LNSLNKEDTSEQVVPVVLVKNWKECAKTRKQPKVYSSKDPAPKPKSSKVDANEVEAL 300  
 Db 241 LNSLNKEDTSEQVVPVVLVKNWKECAKTRKQPKVYSSKDPAPKPKSSKVDANEVEAL 300  
 QY 301 IVKSPKQKWNPSLFKVLKYPGPFYFLMSPFFFAIHDLMFSGPQILKLLIKFVNDTKAPD 360  
 Db 301 IVKSPKQKWNPSLFKVLKYPGPFYFLMSPFFFAIHDLMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTAQLTLVLHGYPHICFVSGMRKINTAVIGAYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVTAQLTLVLHGYPHICFVSGMRKINTAVIGAYRKALVITNSARKSSTV 420  
 QY 421 GEIVNLMSVDAORFMDLATYINMISAPLOVILALYLLWNLGSPVLGAVVWMLMVPYN 480  
 Db 421 GEIVNLMSVDAORFMDLATYINMISAPLOVILALYLLWNLGSPVLGAVVWMLMVPYN 480  
 QY 481 AVYAMKTKTYQVAHMKSKDNRIKLNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540  
 Db 481 AVYAMKTKTYQVAHMKSKDNRIKLNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAYLSAVGTFTVCTPPLVALCTFAVYVTTIDENNILDAQTAFVSLALFNILRFPNLILP 600  
 Db 541 KSAYLSAVGTFTVCTPPLVALCTFAVYVTTIDENNILDAQTAFVSLALFNILRFPNLILP 600  
 QY 601 WYISSIVQASVLSKRLRIFLSHEELEPDSIERPDKGGTNSITVRNATFTWASDPT 660  
 Db 601 WYISSIVQASVLSKRLRIFLSHEELEPDSIERPDKGGTNSITVRNATFTWASDPT 660  
 QY 661 LNGITFSIPEGALVAVVGQVGGKSSLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720  
 Db 661 LNGITFSIPEGALVAVVGQVGGKSSLSALLAEMDKVEGHVAIKGSVAVYVQQAWIQND 720  
 QY 721 SURENIFGCOLLEPYRSVIOACALLPDLEILPSGDRTEIGEKGYNLSGGQKQVSLAR 780  
 Db 721 SURENIFGCOLLEPYRSVIOACALLPDLEILPSGDRTEIGEKGYNLSGGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSANDAHVKGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 Db 781 AVYSNADIYLFDDPLSANDAHVKGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
 QY 841 MSGGKISEMSYQELLARDGAFELRTVYASTEQODAEENGVTGSGPGKEAKQWENG 900  
 Db 841 MSGGKISEMSYQELLARDGAFELRTVYASTEQODAEENGVTGSGPGKEAKQWENG 900  
 QY 901 LVTDSAGKOLQRLSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADRAQTQVVKL 960  
 Db 901 LVTDSAGKOLQRLSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADRAQTQVVKL 960  
 QY 961 SYVDYMKAIGLFISFLSIFLPMCHNVLSALSNYMLSLWTDPIVNGTQHTKVLRSYVG 1020  
 Db 961 SYVDYMKAIGLFISFLSIFLPMCHNVLSALSNYMLSLWTDPIVNGTQHTKVLRSYVG 1020  
 QY 1021 ALGISOGIAVFGYSMAVSGGILASRCLHVDLLHLSRSPMFFERTPSGNLVNRFSEL 1080  
 Db 1021 ALGISOGIAVFGYSMAVSGGILASRCLHVDLLHLSRSPMFFERTPSGNLVNRFSEL 1080  
 QY 1081 DTVDMSIPEVIMKMGSLFNVIACIVILLATPIAIIIPPLGLIYFFVQRYVYASSRQL 1140  
 Db 1081 DTVDMSIPEVIMKMGSLFNVIACIVILLATPIAIIIPPLGLIYFFVQRYVYASSRQL 1140  
 QY 1141 KRLESVSRSPVSHFNETLLGVSIVRAPEEQERFTHQSDLKVDENOKAYYPSIVANRMLA 1200

Db 1141 KRLESVSRSPVSHFNETLLGVSIVRAPEEQERFTHQSDLKVDENOKAYYPSIVANRMLA 1200  
 QY 1201 VRLECVGNICVILFAALFAVISHRSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
 Db 1201 VRLECVGNICVILFAALFAVISHRSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETKEAPWQIOETAPPSSWPQVGRVPEPRNYCLRYREDLDVFLRHINVTINGG 1320  
 Db 1261 VERLKEYSETKEAPWQIOETAPPSSWPQVGRVPEPRNYCLRYREDLDVFLRHINVTINGG 1320  
 QY 1321 EKVGVGTGAGKSSLTGLFRINESARGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 Db 1321 EKVGVGTGAGKSSLTGLFRINESARGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLSRMNLDPFSQYSDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 Db 1381 SGLSRMNLDPFSQYSDEEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 QY 1441 ARALLRKTILVLDATAAVALDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTILVLDATAAVALDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGETOEYCAPSDLLQORGLFYSMKADAGLV 1531  
 Db 1501 DKGETOEYCAPSDLLQORGLFYSMKADAGLV 1531

RESULT 8

ADB20865  
 ID ADB20865 standard; protein; 1531 AA.

XX ADB20865;

XX DT 20-NOV-2003 (first entry)

XX DE MRPI based cancer related protein SEQ ID NO:678.

XX KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;  
 lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;  
 variant allele; multidrug resistance protein 1; MRPI; cytostatic.

XX OS Unidentified.

XX PN WO2003013533-A2.

XX PD 20-FEB-2003.

XX PF 23-JUL-2002; 2002WO-EP008200.

XX PR 23-JUL-2001; 2001EP-00117608.

XX PR 24-MAY-2002; 2002EP-00011710.

XX PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

XX PI Heinrich G, Kerb R;

XX WPI; 2003-354397/33.

XX PT Use of irinotecan or its derivative for preparation of a pharmaceutical  
 composition for treating cancer in a subject having a genome with a  
 variant allele comprising a multidrug resistance protein 1  
 polynucleotide.

XX PS Disclosure; SEQ ID NO 678; 100pp; English.

XX CC The present invention describes a method for the use of irinotecan (I) or  
 its derivative for the preparation of a pharmaceutical composition for  
 treating colorectal, cervical, gastric, lung, ovarian or pancreatic  
 cancer, or malignant glioma in a subject having a genome with a variant  
 allele which comprises a multidrug resistance protein 1 (MRPI)  
 polynucleotide (II). (I) has cytostatic activity. (I) or its derivative  
 can be used for the preparation of a pharmaceutical composition for  
 treating colorectal, cervical, gastric, lung, ovarian or pancreatic



CC cancer, or malignant glioma in a subject, where the subject is a human  
 CC (preferably African or Asian) or a mouse. The present sequence represents  
 CC a sequence which is used in the exemplification of the present invention.  
 XX  
 SQ Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 6; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGCSADGSDPLMDWNTWNTSNPDKFQNTLVVWPCFYLMWACFPFFFLYSRH 60  
 DB 1 MALRGCSADGSDPLMDWNTWNTSNPDKFQNTLVVWPCFYLMWACFPFFFLYSRH 60  
 QY 61 DRGYIOWTPLNKTALGFLWLVWADLFYSPWERSRGIFLAPVLSPTLLGITLLA 120  
 DB 61 DRGYIOWTPLNKTALGFLWLVWADLFYSPWERSRGIFLAPVLSPTLLGITLLA 120  
 QY 121 TFLIQLERRKGQSSGIMLTFLVALVLCALAILRSKIMTALKEDAQVDLFRDITFVYVS 180  
 DB 121 TFLIQLERRKGQSSGIMLTFLVALVLCALAILRSKIMTALKEDAQVDLFRDITFVYVS 180  
 QY 181 LLLIQLVLCFSDRSPLSETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LLLIQLVLCFSDRSPLSETHDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSEOVVPLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL 300  
 DB 241 LWSLNKEDTSEOVVPLVKNWKECAKTRKQPVVYSSKDPAPKESKVDANEVEAL 300  
 QY 301 IVKSPQKWNPSLFVLYKTGPGYFLMSFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 DB 301 IVKSPQKWNPSLFVLYKTGPGYFLMSFFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFVTLVFTACLOTILVHOYPHI CFVSGMRKTAIVGAYRKALVITNSARKSTV 420  
 DB 361 WQGYFVTLVFTACLOTILVHOYPHI CFVSGMRKTAIVGAYRKALVITNSARKSTV 420  
 QY 421 GEIVNLMSVDAQRFMDLATYINNIWAPLOVILALVLLWNLGSPVLGAVVWMLVFPN 480  
 DB 421 GEIVNLMSVDAQRFMDLATYINNIWAPLOVILALVLLWNLGSPVLGAVVWMLVFPN 480  
 QY 481 AVAMKTKTYQVAHMKSDNRKILMNEILNGIKVLKYAWELAFKDKVLAIQEEELVKL 540  
 DB 481 AVAMKTKTYQVAHMKSDNRKILMNEILNGIKVLKYAWELAFKDKVLAIQEEELVKL 540  
 QY 541 KSAYLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFAVSLALFNILRPLNLP 600  
 DB 541 KSAYLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFAVSLALFNILRPLNLP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGNSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGGNSITVRNATFTWARSDDPT 660  
 QY 661 LINGITFSTPEGALVAVGQVCGGKSILLSALLAEMDKVEGHVAKGSVAVYVPOQAWIQND 720  
 DB 661 LINGITFSTPEGALVAVGQVCGGKSILLSALLAEMDKVEGHVAKGSVAVYVPOQAWIQND 720  
 QY 721 SILRENILFGCOLLEPYRYSVIOACALLPDLLEILPSGDRTEIGEGVNLGGQKORVSLAR 780  
 DB 721 SILRENILFGCOLLEPYRYSVIOACALLPDLLEILPSGDRTEIGEGVNLGGQKORVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVIIV 840  
 DB 781 AVYSNADIYLFDDPLSAVDHVGKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVIIV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFALFTYASTEQDABENGVTGVSQPGKEAKQMENG 900  
 DB 841 MSGGKISEMGSYQELLARDGAFALFTYASTEQDABENGVTGVSQPGKEAKQMENG 900  
 QY 901 LVTDPSAGKOLQRLSSSSSYSGDISRHNSFAELQKAEAKKEETWKLMEADKATGQVKL 960  
 DB 901 LVTDPSAGKOLQRLSSSSSYSGDISRHNSFAELQKAEAKKEETWKLMEADKATGQVKL 960

QY 961 SVYWDYMKAIGLFISFLSIFLFCMCHVSNALSNYMLSLWTDPIVNGTOEHTKVRLSVYG 1020  
 DB 961 SVYWDYMKAIGLFISFLSIFLFCMCHVSNALSNYMLSLWTDPIVNGTOEHTKVRLSVYG 1020  
 QY 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILASPMSPFFERTPSGNLVNRFPSKEL 1080  
 DB 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILASPMSPFFERTPSGNLVNRFPSKEL 1080  
 QY 1081 DTVDWSIMEVIVKMFWSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFVYVASSROL 1140  
 DB 1081 DTVDWSIMEVIVKMFWSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFVYVASSROL 1140  
 QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLVKVDENQKAYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLVKVDENQKAYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVSVSYLOVTTYLNWLVMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVISHRSLSAGLVSVSYLOVTTYLNWLVMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNYCLARYREDLDFVLRHINVTINGG 1320  
 DB 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNYCLARYREDLDFVLRHINVTINGG 1320  
 QY 1321 EKVGIIVGTGAGKSLTGLFRINESARBEIIIDGINIAKIGLHDLRPFKTIIPQDPVLF 1380  
 DB 1321 EKVGIIVGTGAGKSLTGLFRINESARBEIIIDGINIAKIGLHDLRPFKTIIPQDPVLF 1380  
 QY 1381 SGSLRMNLDPFSQYSDDEVWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGORQLVCL 1440  
 DB 1381 SGSLRMNLDPFSQYSDDEVWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGORQLVCL 1440  
 QY 1441 ARALLRKTKILVLDATAAVALDLETDLLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ARALLRKTKILVLDATAAVALDLETDLLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIOEGAPSDLLQORGLFYSMAKDAGLV 1531  
 DB 1501 DKGEIOEGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 9

ADB87954  
 ID ADB87954 standard; protein; 1531 AA.

AC ADB87954;

XX ADB87954;  
 DT 04-DEC-2003 (first entry)

XX Human UGT1A1 protein sequence SEQ ID NO:678.

DE irinotecan; cancer; UGT1A1; cytostatic; topoisomerase I inhibitor;  
 KW colorectal cancer; cervical cancer; gastric cancer; lung cancer;  
 KW ovarian cancer; pancreatic cancer; malignant glioma;  
 KW uridine diphosphate glycosyltransferase1 member A1.

OS Homo sapiens.

XX WO2003013536-A2.

FN 20-FEB-2003.

PD 23-JUL-2002; 2002WO-EP008217.

PF 23-JUL-2001; 2001EP-00117608.

PR 24-MAY-2002; 2002EP-00011710.

XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

XX Heinrich G, Kerb R;

XX WPI; 2003-289896/28.



XX Use of irinotecan to treat cancer patient by determining if patient has  
PT variant alleles of UGT1A1 gene, administering increased/decreased amounts  
PT of irinotecan based on increased/decreased levels of UGT1A1 gene product.  
XX  
PS Disclosure; SEQ ID NO 678; 107pp; English.  
XX  
CC The invention relates to the novel use of irinotecan to treat a patient  
CC suffering from cancer. This involves determining if the patient has one  
CC or more variant alleles of the UGT1A1 gene, and if the patient has one  
CC more of such variant alleles, irinotecan is administered in an increased  
CC or decreased amount in comparison to the amount that is administered  
CC without regard to the patient's alleles in the UGT1A1 gene. The invention  
CC has cytostatic activity. A composition of the invention acts as a  
CC topoisomerase I inhibitor. The method is useful for treating a patient,  
CC an animal e.g. mouse or a human, preferably African or Asian, suffering  
CC from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,  
CC pancreatic cancer or malignant glioma. The present sequence is udes in  
CC the exemplification of the invention.  
XX  
SQ Sequence 1531 AA;  
Query Match 78.5%; Score 7860; DB 7; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MALRGFCSADGSDPLMDWNVWNTSNPDTKCFQNTLVVWPCFYLWACFPFYLYLSRH 60  
DB 1 MALRGFCSADGSDPLMDWNVWNTSNPDTKCFQNTLVVWPCFYLWACFPFYLYLSRH 60  
QY 61 DRGYQMTPLNKTALGFLLLIVCWADLFYFWSRSGIFLAPVFLVSPDLLGITLLA 120  
DB 61 DRGYQMTPLNKTALGFLLLIVCWADLFYFWSRSGIFLAPVFLVSPDLLGITLLA 120  
QY 121 TFLIQLERKGVQSSGIMLTFWLVALVCAILRLSKIMTALKEDAQVDLFRDITFYVYFS 180  
DB 121 TFLIQLERKGVQSSGIMLTFWLVALVCAILRLSKIMTALKEDAQVDLFRDITFYVYFS 180  
QY 181 LLLIQLVLSFCSDRSPFLSETIHDNPPCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
DB 181 LLLIQLVLSFCSDRSPFLSETIHDNPPCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
QY 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVVYSSKDPAPKSSKVDANEVEAL 300  
DB 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVVYSSKDPAPKSSKVDANEVEAL 300  
QY 301 IVKSPQKWNPSLFKVLKYPFGYFLMSPFFKAIDHLMFSGPQILKLLIKFVNDTKAPD 360  
DB 301 IVKSPQKWNPSLFKVLKYPFGYFLMSPFFKAIDHLMFSGPQILKLLIKFVNDTKAPD 360  
QY 361 WQGYFVTVLLFVTACLOTLVHLYPHI CFVSGMRKTAIVGAVYRKALVITNSARKSTV 420  
DB 361 WQGYFVTVLLFVTACLOTLVHLYPHI CFVSGMRKTAIVGAVYRKALVITNSARKSTV 420  
QY 421 GEIVNLMSVDAQRFMDLATYINNIWSAPLOVILALYLWNLGSPVLAGVAVVWLMPVN 480  
DB 421 GEIVNLMSVDAQRFMDLATYINNIWSAPLOVILALYLWNLGSPVLAGVAVVWLMPVN 480  
QY 481 AVAMKTKTYQVAHMKSKDNRIKLMEILNGIKVLKLYAWELAFKDKVLAIRQEBELKVLK 540  
DB 481 AVAMKTKTYQVAHMKSKDNRIKLMEILNGIKVLKLYAWELAFKDKVLAIRQEBELKVLK 540  
QY 541 KSAYLSAVCTFTWVCTPFLVALCTFAVYVITDENNILDQAOTAFVSLAFNLIIRPPLNITLP 600  
DB 541 KSAYLSAVCTFTWVCTPFLVALCTFAVYVITDENNILDQAOTAFVSLAFNLIIRPPLNITLP 600  
QY 601 MWISSIVQASVSLKRLIRIFLSHEELEPPDSIERPPVKDGGGTSITVRNATFTWASDDPT 660  
DB 601 MWISSIVQASVSLKRLIRIFLSHEELEPPDSIERPPVKDGGGTSITVRNATFTWASDDPT 660  
QY 661 LINGITFISIEGALVAVVGQVCGCKSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720  
DB 661 LINGITFISIEGALVAVVGQVCGCKSLLSALLAEMDKVEGHVAIKGSVAVYVPPQAWIQND 720

QY 721 SIRENIFGCOLBEEPYRSVIOACALLPDLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
DB 721 SIRENIFGCOLBEEPYRSVIOACALLPDLEILPSGDRTEIGEKGVLNLSGGQKQVSLAR 780  
QY 781 AVYSNADYILFDDPLSAVDHVGKHIPENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
DB 781 AVYSNADYILFDDPLSAVDHVGKHIPENVIGPKGMLKNKTRILVTHSMSYLPQVDVIV 840  
QY 841 MSGGKISMGVSQELLARDGAFABFLRYASTEQDQAEENGVTGVSFGPKAKOMENGM 900  
DB 841 MSGGKISMGVSQELLARDGAFABFLRYASTEQDQAEENGVTGVSFGPKAKOMENGM 900  
QY 901 LVTSAGLQRLQRLSSSSSSSGDISRHNSTAELOKAEAKKEETWKLMEADKAQTCQVKL 960  
DB 901 LVTSAGLQRLQRLSSSSSSSGDISRHNSTAELOKAEAKKEETWKLMEADKAQTCQVKL 960  
QY 961 SVYDYMKAIGLFTSIFLIFPMCNHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
DB 961 SVYDYMKAIGLFTSIFLIFPMCNHVSALASNYWLSLWTDPIVNGTQEHTKVRLSVYG 1020  
QY 1021 ALGISQGIATVFGYSMAVISGIGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
DB 1021 ALGISQGIATVFGYSMAVISGIGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSEL 1080  
QY 1081 DTVDSMIEVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQREYFVASSRQL 1140  
DB 1081 DTVDSMIEVIMKFMGSLFNVIACIVILLATPIAAIIIPPLGLIYFFVQREYFVASSRQL 1140  
QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
DB 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
QY 1201 VRLECVGNICVILFAALFAVIRHSHSAGLVGSVSYSLQVTTYLANLWVMSSEMETNIVA 1260  
DB 1201 VRLECVGNICVILFAALFAVIRHSHSAGLVGSVSYSLQVTTYLANLWVMSSEMETNIVA 1260  
QY 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVEFRNYCLRYREDLDVFLRHINVTINGG 1320  
DB 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVEFRNYCLRYREDLDVFLRHINVTINGG 1320  
QY 1321 EKVGIIVGTGAGKSLTGLFRINESARGEIIDGINIAKIGLHDLRFPKTIIPQDPVLF 1380  
DB 1321 EKVGIIVGTGAGKSLTGLFRINESARGEIIDGINIAKIGLHDLRFPKTIIPQDPVLF 1380  
QY 1381 SCSLRMNLDPSOYSDDEEVTSLAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
DB 1381 SCSLRMNLDPSOYSDDEEVTSLAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
QY 1441 ARALLRKTILVDEATAVDETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
DB 1441 ARALLRKTILVDEATAVDETDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
QY 1501 DKGETQYGAESDILLQORGLFYFMAKADGLV 1531  
DB 1501 DKGETQYGAESDILLQORGLFYFMAKADGLV 1531

RESULT 10  
ADB96937

ID ADB96937 standard; protein; 1531 AA.

XX ADB96937;

XX 04-DEC-2003 (first entry)

XX Human MDR1 related protein sequence SEQ ID NO:678.

XX irinotecan; colorectal cancer; cervical cancer; gastric cancer;  
KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;  
KW multidrug resistance 1; MDR1; cytostatic; human; CYP3A5; MRP1; MDR1;  
XX TOPI.

OS Homo sapiens.  
 PN WC2003013537-A2.  
 PD 20-FEB-2003.  
 XX 23-JUL-2002; 2002WO-EP008218.  
 XX 23-JUL-2001; 2001EP-00117608.  
 PR 24-MAY-2002; 2002EP-00011710.  
 XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
 PA Heinrich G, Kerb R;  
 XX WPI; 2003-268145/26.  
 XX  
 PT New use of irinotecan for preparation of pharmaceutical compositions for  
 PT treating cancer in subject having genome with variant allele comprising  
 PT multidrug resistance 1 polynucleotide.  
 XX  
 PS Disclosure; SEQ ID NO 678; 130pp; English.  
 XX  
 CC The invention relates to the novel use of irinotecan or its derivative  
 CC for the preparation of pharmaceutical compositions for treating  
 CC colorectal, cervical, gastric, lung, ovarian or pancreatic cancer, or  
 CC malignant glioma in a subject having a genome with a variant allele which  
 CC comprises a multidrug resistance 1 (MDR1) polynucleotide. A composition  
 CC of the invention has cytostatic activity. The invention is useful for the  
 CC preparation of pharmaceutical compositions for treating colorectal,  
 CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant  
 CC glioma in a subject (preferably human, more preferably African or Asian)  
 CC or a mouse. The present sequence is used in the exemplification of the  
 CC invention.  
 XX  
 SQ Sequence 1531 AA;  
 Query Match 78.5%; Score 7860; DB 7; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRFGCSADGDPDLDWNVNTWNTSNPDFTKCFQNTVLVWVPCFVLMACFPFFFLYLSRH 60  
 DB 1 MALRFGCSADGDPDLDWNVNTWNTSNPDFTKCFQNTVLVWVPCFVLMACFPFFFLYLSRH 60  
 QY 61 DRGYIQTPLNKTALGFLWVWADLFYSFWSRSRGIFLAPVFLVSPFLGTTLLA 120  
 DB 61 DRGYIQTPLNKTALGFLWVWADLFYSFWSRSRGIFLAPVFLVSPFLGTTLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 QY 181 LLLIQLVLSFSDRSPLETHDNPCEPESASFLSRITFWITGLIVRGYRQPLGSD 240  
 DB 181 LLLIQLVLSFSDRSPLETHDNPCEPESASFLSRITFWITGLIVRGYRQPLGSD 240  
 QY 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPVKVYSSKDPAPKSSKVDANEVEAL 300  
 DB 241 LWSLNKEDTSEQVPLVKNWKECAKTRKQPVKVYSSKDPAPKSSKVDANEVEAL 300  
 QY 301 IVKSPQKWNPSLFKVLVYKTPGFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 DB 301 IVKSPQKWNPSLFKVLVYKTPGFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTAQLVHLVHCFVSGMRKTAIVGAVYKALVITNSARKSSTV 420  
 DB 361 WQGYFYTVLLFVTAQLVHLVHCFVSGMRKTAIVGAVYKALVITNSARKSSTV 420  
 QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLMLNGLPSVLGAVVMVMPVN 480  
 DB 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLMLNGLPSVLGAVVMVMPVN 480

QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540  
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAQAFVSLALFNLIRPLNLLP 600  
 DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNILDQAQAFVSLALFNLIRPLNLLP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGSGTNSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGSGTNSITVRNATFTWARSDDPT 660  
 QY 661 LINGITFISPEGALVAVVGVCGKSSLLSALLAEWMDKVEGHVAKGSAVYVPOQAWTQND 720  
 DB 661 LINGITFISPEGALVAVVGVCGKSSLLSALLAEWMDKVEGHVAKGSAVYVPOQAWTQND 720  
 QY 721 SLRENILFGCQLEBPYYSRVIQACALLPDLEILPSGDRTEIGEGVNLSSGQKORVSLAR 780  
 DB 721 SLRENILFGCQLEBPYYSRVIQACALLPDLEILPSGDRTEIGEGVNLSSGQKORVSLAR 780  
 QY 781 AVYSNADIYLPDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIIV 840  
 DB 781 AVYSNADIYLPDDPLSAVDHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVIIV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEOQDAENGVTGVSQKGAKEOMENGM 900  
 DB 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEOQDAENGVTGVSQKGAKEOMENGM 900  
 QY 901 LVTDGAKQLQRLQSSSSSYSGDTSRHNHSTAEILQKAEKEETWKLMEADKAQTGVKL 960  
 DB 901 LVTDGAKQLQRLQSSSSSYSGDTSRHNHSTAEILQKAEKEETWKLMEADKAQTGVKL 960  
 QY 961 SVYDYNKALGLFISFLSIFLFCNHNVSALASYNWLSLWTDTPVNGTQSHTKVRLSVYG 1020  
 DB 961 SVYDYNKALGLFISFLSIFLFCNHNVSALASYNWLSLWTDTPVNGTQSHTKVRLSVYG 1020  
 QY 1021 ALGISQIAGVFGSMAYSIGIILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEL 1080  
 DB 1021 ALGISQIAGVFGSMAYSIGIILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEL 1080  
 QY 1081 DTVDSPITPEVVKPMFMSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 DB 1081 DTVDSPITPEVVKPMFMSLNFVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 QY 1141 KRLESVRSRPVYSHENETLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 DB 1141 KRLESVRSRPVYSHENETLGVSVIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVIRSHLSLGLSVSYSLQVTTVNLNMLVRMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVIRSHLSLGLSVSYSLQVTTVNLNMLVRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKAPWQIQETAPPSPWPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320  
 DB 1261 VERLKEYSETEKAPWQIQETAPPSPWPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320  
 QY 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 DB 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGSLRMNLDPFQSYDEEYVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL 1440  
 DB 1381 SGSLRMNLDPFQSYDEEYVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL 1440  
 QY 1441 ARALLRKTLLVLDEATAVDLETDLLIQSTIRTOQEDCTVLTIAHRLNTIMDYTRVIL 1500  
 DB 1441 ARALLRKTLLVLDEATAVDLETDLLIQSTIRTOQEDCTVLTIAHRLNTIMDYTRVIL 1500  
 QY 1501 DKGEIYEGAPSDLLQORGLFYSMKADGLV 1531  
 DB 1501 DKGEIYEGAPSDLLQORGLFYSMKADGLV 1531

RESULT 11  
AD892128  
ID ADB92128 standard; protein; 1531 AA.  
XX AC ADB92128;  
XX DT 04-DEC-2003 (first entry)  
XX DE Human MDR1 related protein sequence SEQ ID NO:678.  
XX KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;  
XX KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;  
XX KW multidrug resistance 1; MDR1; cytostatic; human; UGT1A1; MRP1; TOP1.  
XX OS Homo sapiens.  
XX PN WO2003013535-A2.  
XX PD 20-FEB-2003.  
XX PF 23-JUL-2002; 2002WO-EP008220.  
XX PR 23-JUL-2001; 2001EP-00117608.  
XX PR 24-MAY-2002; 2002EP-00011710.  
XX PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
XX PI Heinrich G, Kerb R;  
XX WPI; 2003-342400/32.  
XX  
PT New use of irinotecan for preparation of pharmaceutical compositions for  
PT treating cancer in subject having genome with variant allele comprising  
PT multidrug resistance 1 polynucleotide.  
XX  
PS Disclosure; SEQ ID NO 678; 104pp; English.  
XX  
CC The invention relates to a novel use of irinotecan or its derivative for  
CC the preparation of a pharmaceutical composition for treating colorectal,  
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant  
CC glioma in a subject having a genome with a variant allele which comprises  
CC a multidrug resistance 1 (MDR1) polynucleotide. A composition of the  
CC invention has cytostatic activity. The present sequence is used in the  
CC exemplification of the invention.  
XX  
SQ Sequence 1531 AA;  
Query Match 78.5%; Score 7860; DB 7; Length 1531;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MALRGFCSADGSDPLMDWNVNTNSNPDTKCFQNTVLVWVPCFYLWACFPFYLGRH 60  
DB 1 MALRGFCSADGSDPLMDWNVNTNSNPDTKCFQNTVLVWVPCFYLWACFPFYLGRH 60  
QY 61 DRGYQMTPLNKTALGFLLLIWCWADLFYFWSRSGIFLAPFLVSPDLLGITLLA 120  
DB 61 DRGYQMTPLNKTALGFLLLIWCWADLFYFWSRSGIFLAPFLVSPDLLGITLLA 120  
QY 121 TFLIQLERRKGVSQSGIMLTFWLVVALVCAALILRSKIMTALKEDAQVDLFRDITYVYFS 180  
DB 121 TFLIQLERRKGVSQSGIMLTFWLVVALVCAALILRSKIMTALKEDAQVDLFRDITYVYFS 180  
QY 181 LLLIQLVLSGSDRPLSETHDNPCCPESASFLSRITFWITGLIVRGVROPLEGSD 240  
DB 181 LLLIQLVLSGSDRPLSETHDNPCCPESASFLSRITFWITGLIVRGVROPLEGSD 240  
QY 241 LMSLNKEDTSEQVPLVKNWKECAKTRKQPVVYSSKDPAPKSSKVDANEVEAL 300  
DB 241 LMSLNKEDTSEQVPLVKNWKECAKTRKQPVVYSSKDPAPKSSKVDANEVEAL 300  
QY 301 IVKSPQKWNPSLFKVLKTFPGYFLMSFFFAKIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
DB 301 IVKSPQKWNPSLFKVLKTFPGYFLMSFFFAKIHDLMMFSGPQILKLLIKFVNDTKAPD 360

DB 301 IVKSPQKWNPSLFKVLKTFPGYFLMSFFFAKIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
QY 361 WQGYFTVLLFVTACLOTLVLHOYPHICFVSGMRKTKTAVIGAVYKALVITNSARKSSTV 420  
DB 361 WQGYFTVLLFVTACLOTLVLHOYPHICFVSGMRKTKTAVIGAVYKALVITNSARKSSTV 420  
QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLLNLGSPSLAGVAVMLMVPVN 480  
DB 421 GEIVNLSVDAQRFMDLATYINMIWSAPLQVILALYLLNLGSPSLAGVAVMLMVPVN 480  
QY 481 AVMAKTKTYQVAHMKSKDNRIKLNNELINGIKVLKYAWELAFKDKVLAIHQBELKVLK 540  
DB 481 AVMAKTKTYQVAHMKSKDNRIKLNNELINGIKVLKYAWELAFKDKVLAIHQBELKVLK 540  
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVVTVDENNILDAQTAFAVSLALFNILFPLNILP 600  
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVVTVDENNILDAQTAFAVSLALFNILFPLNILP 600  
QY 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRRPVKGGGNSITVRNATFTWASDDPPT 660  
DB 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRRPVKGGGNSITVRNATFTWASDDPPT 660  
QY 661 LINGITFSIPEGALVAVGVGCGKSSLSALLAEMDKVEGHVAKIGSVAYVPOQAWI QND 720  
DB 661 LINGITFSIPEGALVAVGVGCGKSSLSALLAEMDKVEGHVAKIGSVAYVPOQAWI QND 720  
QY 721 SLRENILFGCOLLEPYRSVIOACALLPDLLEIPSGDRTEIGEKGNSLGGQKQVSLAR 780  
DB 721 SLRENILFGCOLLEPYRSVIOACALLPDLLEIPSGDRTEIGEKGNSLGGQKQVSLAR 780  
QY 781 AVYSNADTYLFDPLSADVAHVGHIFENVIGPKMGLKNKTRILVTHSMYLPQVDVIV 840  
DB 781 AVYSNADTYLFDPLSADVAHVGHIFENVIGPKMGLKNKTRILVTHSMYLPQVDVIV 840  
QY 841 MSGGKISEMSYQELLARDGAFAPFLRTYASTEQODAEENGVTGSGPGEAKOMENGM 900  
DB 841 MSGGKISEMSYQELLARDGAFAPFLRTYASTEQODAEENGVTGSGPGEAKOMENGM 900  
QY 901 LVTDGAGLQORQLSSSSSYSGDISRHNSHNSAELOKAEKKEETWKLMEADKAQGOVKL 960  
DB 901 LVTDGAGLQORQLSSSSSYSGDISRHNSHNSAELOKAEKKEETWKLMEADKAQGOVKL 960  
QY 961 SVYDYMKAIGLIFISFLSIFLFCMCHVSNALSNYLSLWTDPIVNGTQHTKVRLSVYG 1020  
DB 961 SVYDYMKAIGLIFISFLSIFLFCMCHVSNALSNYLSLWTDPIVNGTQHTKVRLSVYG 1020  
QY 1021 ALGISQGIATVFGYSNAVSIGGILASRCLHVDLLHLSILRSPMSFFERTSGNLVNRFSKEL 1080  
DB 1021 ALGISQGIATVFGYSNAVSIGGILASRCLHVDLLHLSILRSPMSFFERTSGNLVNRFSKEL 1080  
QY 1081 DTVDSMIPEVIMKFMGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSROL 1140  
DB 1081 DTVDSMIPEVIMKFMGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSROL 1140  
QY 1141 KRLESVRSRSPVYSHFNETHLLGVSVIRAFEBEGRFTHQSDLKVDENOKAYYPSIVANRWLA 1200  
DB 1141 KRLESVRSRSPVYSHFNETHLLGVSVIRAFEBEGRFTHQSDLKVDENOKAYYPSIVANRWLA 1200  
QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGVLSVSYLQVTTYLNLVMSSEMETNIVA 1260  
DB 1201 VRLECVGNCIVLFAALFAVISRHSLSAGVLSVSYLQVTTYLNLVMSSEMETNIVA 1260  
QY 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVEFRNYCLRYREDLDFVLRHINTVINGG 1320  
DB 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVEFRNYCLRYREDLDFVLRHINTVINGG 1320  
QY 1321 EKVGVGRTGAGKSSLTGLFRINESAEGEIIDGINIAKIGHDLRPFKTIIPDPVLF 1380  
DB 1321 EKVGVGRTGAGKSSLTGLFRINESAEGEIIDGINIAKIGHDLRPFKTIIPDPVLF 1380  
QY 1381 SGLSRMNLDPFSQVSDSEBWTSLAHLKDFVSALPDKLDHECAGGENLSVQRLVCL 1440  
DB 1381 SGLSRMNLDPFSQVSDSEBWTSLAHLKDFVSALPDKLDHECAGGENLSVQRLVCL 1440

QY 1441 ARALLRKTILVLDEATAAVDLEDDLIQSTIOTPEDECTVLIAHRLNTIMDYTRVIVL 1500  
 DB |||||||  
 QY 1441 ARALLRKTILVLDEATAAVDLEDDLIQSTIOTPEDECTVLIAHRLNTIMDYTRVIVL 1500  
 DB |||||||  
 QY 1501 DKGEIQEYGAPSDLLQORGLFYSMAKDAGLV 1531  
 DB |||||||  
 QY 1501 DKGEIQEYGAPSDLLQORGLFYSMAKDAGLV 1531  
 DB |||||||

RESULT 12  
 ADD44764  
 ID ADD44764 standard; protein; 1531 AA.  
 XX  
 AC ADD44764;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human Protein P33527, SEQ ID NO 10193.  
 XX  
 KW Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 14-AUG-2002; 2002WO-US025765.  
 XX  
 PR 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX  
 XX (GEO) GEN HOSPITAL CORP.  
 PA (FARB) BAYER AG.  
 XX  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX  
 XX WPI; 2003-268312/26.  
 DR GENBANK; P33527.  
 XX  
 XX  
 PT New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.  
 PS  
 PS Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:

CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 1531 AA;  
 Query Match 78.5%; Score 7860; DB 7; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGFCSDGSDPLMDNVTWNTSNPDFTKCFQNTLVWVPCFYLWACFPFYLYLSRH 60  
 DB |||||||  
 QY 1 MALRGFCSDGSDPLMDNVTWNTSNPDFTKCFQNTLVWVPCFYLWACFPFYLYLSRH 60  
 DB |||||||  
 QY 61 DRGYIQMTPLNKTALGFLWIVCWADLFYSFWSRSGIFLAPVFLVSPITLLGTTLLA 120  
 DB |||||||  
 QY 61 DRGYIQMTPLNKTALGFLWIVCWADLFYSFWSRSGIFLAPVFLVSPITLLGTTLLA 120  
 DB |||||||  
 QY 121 TFLIQLEERRKGVQSSGIMLTFLWVALVCALAILRSKIMTALKEDAQVDFRDITFYVYFS 180  
 DB |||||||  
 QY 121 TFLIQLEERRKGVQSSGIMLTFLWVALVCALAILRSKIMTALKEDAQVDFRDITFYVYFS 180  
 DB |||||||  
 QY 181 LLLIQVLSCFSDRSPLESETIHDNPNCPSSASPLSRITFWWITGLIVRGVROPLEGSD 240  
 DB |||||||  
 QY 181 LLLIQVLSCFSDRSPLESETIHDNPNCPSSASPLSRITFWWITGLIVRGVROPLEGSD 240  
 DB |||||||  
 QY 241 LWSLNKEDTSQVVPVVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANEVEAL 300  
 DB |||||||  
 QY 241 LWSLNKEDTSQVVPVVLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANEVEAL 300  
 DB |||||||  
 QY 301 IVKSPQKEWNPFLFKVLYKTFPGPYFLMSFFPKAIDHLMFSGPQILKLIKFVNDTKAPD 360  
 DB |||||||  
 QY 301 IVKSPQKEWNPFLFKVLYKTFPGPYFLMSFFPKAIDHLMFSGPQILKLIKFVNDTKAPD 360  
 DB |||||||  
 QY 361 WQGYFYTVLLFVTACLOTLVHLQYFHICFVSGMRITKAVIGAVYRKALVITNSARKSTV 420  
 DB |||||||  
 QY 361 WQGYFYTVLLFVTACLOTLVHLQYFHICFVSGMRITKAVIGAVYRKALVITNSARKSTV 420  
 DB |||||||  
 QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLNMLNGLSPVLGAVVWVMPVN 480  
 DB |||||||  
 QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLNMLNGLSPVLGAVVWVMPVN 480  
 DB |||||||  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNBIKLVYAWELAPKDKVLAIRQELKVLK 540  
 DB |||||||  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNBIKLVYAWELAPKDKVLAIRQELKVLK 540  
 DB |||||||  
 QY 541 KSAVLSAVGTFTWVCTPFLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRPLNLP 600  
 DB |||||||  
 QY 541 KSAVLSAVGTFTWVCTPFLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRPLNLP 600  
 DB |||||||  
 QY 601 MVISSIVOASVSLKRLRIFLUSHELEPDSIERRPVKGGGTNSITVRNATFTWARSDDPT 660  
 DB |||||||  
 QY 601 MVISSIVOASVSLKRLRIFLUSHELEPDSIERRPVKGGGTNSITVRNATFTWARSDDPT 660  
 DB |||||||  
 QY 661 LMGITFIPREGALVAVVGVQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWTQND 720  
 DB |||||||  
 QY 661 LMGITFIPREGALVAVVGVQVCGCKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWTQND 720  
 DB |||||||  
 QY 721 SLRENILFGQLBEPYRVSUIQACALLPDLEILPSGDRTEIGKGVNLSGGQKQVSLAR 780  
 DB |||||||  
 QY 721 SLRENILFGQLBEPYRVSUIQACALLPDLEILPSGDRTEIGKGVNLSGGQKQVSLAR 780  
 DB |||||||  
 QY 781 AVYSNADIYLFDDPLSAVDARVGHKHFENVIGPKMLKNKTRILVTHSMYSYLPQVDVII 840  
 DB |||||||  
 QY 781 AVYSNADIYLFDDPLSAVDARVGHKHFENVIGPKMLKNKTRILVTHSMYSYLPQVDVII 840  
 DB |||||||  
 QY 841 MSGGKISEMSYQELLARDGAFAEFLRTYASTEQEQAENGVTGSGPCKEAKQEMNGM 900  
 DB |||||||  
 QY 841 MSGGKISEMSYQELLARDGAFAEFLRTYASTEQEQAENGVTGSGPCKEAKQEMNGM 900  
 DB |||||||  
 QY 901 LVTDTSAGKQLOROLSSSSSYSGDISRHNSHTAELQKAEAKKETWKLMEADKATQGVKL 960  
 DB |||||||  
 QY 901 LVTDTSAGKQLOROLSSSSSYSGDISRHNSHTAELQKAEAKKETWKLMEADKATQGVKL 960  
 DB |||||||

QY 961 SYVDYMKAIIGLIFSLIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTVLSVYG 1020  
 DB 961 SYVDYMKAIIGLIFSLIFLPMCNHVSALASNYWLSLWTDPIVNGTQEHKTVLSVYG 1020  
 QY 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILSPMSFFERTSGNLVNRFSKEL 1080  
 DB 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILSPMSFFERTSGNLVNRFSKEL 1080  
 QY 1081 DTVDSMIPEVIMKMGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 DB 1081 DTVDSMIPEVIMKMGSLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 QY 1141 KELESVSRSPVYSHFNETLLGVSIVRAPEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
 DB 1141 KELESVSRSPVYSHFNETLLGVSIVRAPEEQERFIHQSDLKVDENQKAYPSIVANRWLA 1200  
 QY 1201 VRLCEVGNICIVLFAALFAVISHSHLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
 DB 1201 VRLCEVGNICIVLFAALFAVISHSHLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKEAPWQIQTAPPSSWPQVGRVFRNCLRYREDLDFVLRHINVTINGG 1320  
 DB 1261 VERLKEYSETEKEAPWQIQTAPPSSWPQVGRVFRNCLRYREDLDFVLRHINVTINGG 1320  
 QY 1321 EKVIGVGRGACKSSITLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 DB 1321 EKVIGVGRGACKSSITLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLSRMNLDPFQSYDEEVTWTSLELAHLKDFVSALPKLDHECARGGENLSVGQRQLVCL 1440  
 DB 1381 SGLSRMNLDPFQSYDEEVTWTSLELAHLKDFVSALPKLDHECARGGENLSVGQRQLVCL 1440  
 QY 1441 APALARKTKILVLDATAVLDLTDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 APALARKTKILVLDATAVLDLTDLLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIOEYCAPSLLQORGLFYSMADAGLV 1531  
 DB 1501 DKGEIOEYCAPSLLQORGLFYSMADAGLV 1531  
 RESULT 13  
 ADN97111  
 ID ADN97111 standard; protein; 1531 AA.  
 AC ADN97111;  
 XX ADN97111;  
 DT 01-JUL-2004 (first entry)  
 DE MRP1 protein.  
 DE MRP1 protein.  
 KW multiple drug resistance protein; MRP; Drosophila melanogaster;  
 KW Anopheles gambiae; insecticide.  
 XX Homo sapiens.  
 XX WO2004029088-A2.  
 XX 08-APR-2004.  
 XX 25-SEP-2003; 2003WO-EP012400.  
 XX 26-SEP-2002; 2002US-0413469P.  
 XX (INSP ) INST PASTEUR.  
 PA (CNRS ) CENT NAT RECH SCI.  
 XX Roth CW, Brey PT, Holm I, Graillies M, Rzhetsky A;  
 PI WPI; 2004-305150/28.  
 DR New polynucleotide sequence encoding multiple drug resistance proteins  
 PT

PT from Drosophila melanogaster or Anopheles gambiae, useful in developing  
 effective insecticides.  
 XX Claim 1; SEQ ID NO 6; 58pp; English.  
 CC The present invention relates to a purified polynucleotide or its  
 fragment and comprises a sequence encoding multiple drug resistance  
 proteins (MRPs) from Drosophila melanogaster or Anopheles gambiae. The  
 polynucleotide is useful in developing effective insecticides. The  
 present sequence represents human MRP1.  
 XX Sequence 1531 AA;  
 SQ  
 Query Match 78.5%; Score 7860; DB 8; Length 1531;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGFCSADSGDLMDWNTWNTSNPDFTKCFQNTVLVWVPCFVLMACFPFFYFLYLSRH 60  
 DB 1 MALRGFCSADSGDLMDWNTWNTSNPDFTKCFQNTVLVWVPCFVLMACFPFFYFLYLSRH 60  
 QY 61 DRGIQMTPLNKTALGFLWIVCWADLFYSFWERSRGIIFLAPVFLVSPILLGTTLLA 120  
 DB 61 DRGIQMTPLNKTALGFLWIVCWADLFYSFWERSRGIIFLAPVFLVSPILLGTTLLA 120  
 QY 121 TFLQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDIITYVYVS 180  
 DB 121 TFLQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDIITYVYVS 180  
 QY 181 LLLQLVLSGSDRSPLESETIHDNPNCPESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LLLQLVLSGSDRSPLESETIHDNPNCPESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKOPKVYVSSKDPAPQPKESKVDANEVEAL 300  
 DB 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKOPKVYVSSKDPAPQPKESKVDANEVEAL 300  
 QY 301 IVKSPQKEMNPSLFKVLKTEGYPFLMSFFPKAIHDLMMFSGPQILKLIKFVNDTKAPD 360  
 DB 301 IVKSPQKEMNPSLFKVLKTEGYPFLMSFFPKAIHDLMMFSGPQILKLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTACLOTILVHOFHICFVSGMIRKTAIVIGAVYRKALVITNSARKSTV 420  
 DB 361 WQGYFYTVLLFVTACLOTILVHOFHICFVSGMIRKTAIVIGAVYRKALVITNSARKSTV 420  
 QY 421 GEIVNLMSVDAQRFMDLATYINMLWSAPLOVILALYLNLGLPSVLGAVVWMLVNPV 480  
 DB 421 GEIVNLMSVDAQRFMDLATYINMLWSAPLOVILALYLNLGLPSVLGAVVWMLVNPV 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKDKVLAIHQBELKVLK 540  
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKDKVLAIHQBELKVLK 540  
 QY 541 KSAVLSAVGTWCTTVPFLVALCTFAVVTVDENNILDAQTAFVSLALFNILRFPNLILP 600  
 DB 541 KSAVLSAVGTWCTTVPFLVALCTFAVVTVDENNILDAQTAFVSLALFNILRFPNLILP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERPPVKDGGTNSITVRNATFTWARSDDPT 660  
 DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERPPVKDGGTNSITVRNATFTWARSDDPT 660  
 QY 661 LMGITFSIPEGALVAVVGVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPQQAQIWD 720  
 DB 661 LMGITFSIPEGALVAVVGVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPQQAQIWD 720  
 QY 721 SLRENILFGCOLBEPYRSVIOACALLPDLBILPSGDRTEIGEKVNLSSGQKQVSLAR 780  
 DB 721 SLRENILFGCOLBEPYRSVIOACALLPDLBILPSGDRTEIGEKVNLSSGQKQVSLAR 780  
 QY 781 AVYSNADIYLPDDPLSAVDHVGKHIPENVIGPKMKNKTRILVTHSMYSYLPQVDVIV 840  
 DB 781 AVYSNADIYLPDDPLSAVDHVGKHIPENVIGPKMKNKTRILVTHSMYSYLPQVDVIV 840

841 MSGGKISEMSYQELLARDGAFELRTYASTQEODAEENGVTGYSGPCKEAKOMENG 900  
 841 MSGGKISEMSYQELLARDGAFELRTYASTQEODAEENGVTGYSGPCKEAKOMENG 900  
 901 LVTDGAKQIQRLSSSSSYSGDISRHNSHSTABLOKAEAKKEETWKLMEADKAQTQGVKL 960  
 901 LVTDGAKQIQRLSSSSSYSGDISRHNSHSTABLOKAEAKKEETWKLMEADKAQTQGVKL 960  
 961 SVYWDYKATGCLPISFLSIFLFCNHNVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 961 SVYWDYKATGCLPISFLSIFLFCNHNVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYG 1020  
 1021 ALGISOCIAVFGYSMAVSGIGILASCLHVDLHLSILRSPMSFFERTPSGNLVNRSKEL 1080  
 1021 ALGISOCIAVFGYSMAVSGIGILASCLHVDLHLSILRSPMSFFERTPSGNLVNRSKEL 1080  
 1081 DTVDSDIPEVKMFMSLGNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSRQL 1140  
 1081 DTVDSDIPEVKMFMSLGNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFVASSRQL 1140  
 1141 KRLESVSRSPVYSHNETLLGVSVIRAFBEQERFIHQSDLKVDENQKAYPYSIVANRWLA 1200  
 1141 KRLESVSRSPVYSHNETLLGVSVIRAFBEQERFIHQSDLKVDENQKAYPYSIVANRWLA 1200  
 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETNIVA 1260  
 1261 VERLKEYSTETEAPQIQTETAPPSSWPQVGRVFRNRYCLRYREDLDFVLRHINVTINGG 1320  
 1261 VERLKEYSTETEAPQIQTETAPPSSWPQVGRVFRNRYCLRYREDLDFVLRHINVTINGG 1320  
 1321 EKVIGVGRGAGKSSLTGLFRINSAGEIIIDGINIAKIGHDLREKTIIPQDPVLF 1380  
 1321 EKVIGVGRGAGKSSLTGLFRINSAGEIIIDGINIAKIGHDLREKTIIPQDPVLF 1380  
 1381 SGLRNLDPFQYSDDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 1381 SGLRNLDPFQYSDDEEVTWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
 1441 ARALARKTILVLDATAAVDLETDLQSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 1441 ARALARKTILVLDATAAVDLETDLQSTIRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 1501 DKGEIQEYGAPELDLQQRGLFYSMADAGLV 1531  
 1501 DKGEIQEYGAPELDLQQRGLFYSMADAGLV 1531

RESULT 14  
 ADN95929  
 ADN95929 standard; protein; 1530 AA.  
 ADN95929;  
 01-JUL-2004 (first entry)  
 Human BEC/LEC-related protein sequence SeqID853.  
 growth; differentiation; blood endothelial cell; BEC;  
 lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGF-3;  
 lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;  
 vasotrophic; antiinflammatory; gene therapy; endothelial cell disorder;  
 inflammatory disease; cancer metastasis; lymphatic system; human.  
 Homo sapiens.  
 WO2003080640-A1.  
 02-OCT-2003.  
 07-MAR-2003; 2003WO-US006900.

PR 07-MAR-2002; 2002US-0363019P.  
 XX (LUDW-) LUDWIG INST CANCER RES.  
 PA (LICN) LICENTIA LTD.  
 XX Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;  
 DR WPI; 2003-876899/81.  
 DR N-P8DB; ADN95930.  
 PS Example 1; SEQ ID NO 853; 176pp; English.  
 XX This invention relates to a method of differentially modulating the  
 CC growth or differentiation of blood endothelial cells (BEC) or lymphatic  
 CC endothelial cells (LEC) comprises contacting endothelial cells with a  
 CC composition comprising an agent that differentially modulates blood or  
 CC lymphatic endothelial cells. Treating hereditary lymphoedema comprises  
 CC identifying a human subject with lymphoedema and with a mutation in at  
 CC least one allele of a gene encoding a LEC protein, where the mutation  
 CC correlates with lymphoedema in human subjects, and with the proviso that  
 CC the LEC protein is not VEGF-3; and administering to the subject a  
 CC composition comprising a lymphatic growth agent selected from VEGF-C or  
 CC VEGF-D polypeptides and polynucleotides. The invention may be useful for  
 CC the development of compounds with an antiangiogenic, cytostatic,  
 CC vasotropic or antiinflammatory activity or for gene therapy. The method  
 CC is useful in modulating the growth or differentiation of blood  
 CC endothelial cells or lymphatic endothelial cells, in treating hereditary  
 CC lymphoedema, in screening for an endothelial cell disorder or  
 CC predisposition to the disorder or in monitoring the efficacy or toxicity  
 CC of a drug on endothelial cells. The agent is useful in manufacturing a  
 CC medicament for the differential modulation of blood vessel endothelial  
 CC cell or lymphatic vessel endothelial cell growth or differentiation. The  
 CC lymphatic growth agent may also be used in manufacturing a medicament for  
 CC the treatment of hereditary lymphoedema resulting from a mutation in a  
 CC LEC gene or of other diseases involving the lymphatic vessels, such as  
 CC various inflammatory diseases and cancer metastasis via the lymphatic  
 CC system. The present sequence is that of a human LEC/BEC differentially  
 CC expressed protein which is related to the method of the invention. Note:  
 CC This sequence does not appear in the specification but was obtained by  
 CC the indexer using the source data given in table 14 of the specification.  
 XX Sequence 1530 AA;

Query Match 78.4%; Score 7856; DB 7; Length 1530;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MALRGFCSADGSDPLWDNVTWNTSNPDFTKCFONTVLVWVPCFYLWACPPFFLYLSRH 60  
 DB 1 MALRGFCSADGSDPLWDNVTWNTSNPDFTKCFONTVLVWVPCFYLWACPPFFLYLSRH 60  
 QY 61 DRGIQMTPLNKTALGFLLLWVWADLFYSFWSRGIPLAPVFLVSPITLLGTTLLA 120  
 DB 61 DRGIQMTPLNKTALGFLLLWVWADLFYSFWSRGIPLAPVFLVSPITLLGTTLLA 120  
 QY 121 TFLIQLERRRGVSSGIMLTFMLVALVCAIILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TFLIQLERRRGVSSGIMLTFMLVALVCAIILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 QY 181 LLLIQLVLSGSDRSPISFSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LLLIQLVLSGSDRSPISFSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRQPKVYVSSKDPAPQKSSKVDANEEVEAL 300  
 DB 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRQPKVYVSSKDPAPQKSSKVDANEEVEAL 300  
 QY 301 IVKSPQKWNPSLFKVLKTYKTFGYFLMSPPFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 DB 301 IVKSPQKWNPSLFKVLKTYKTFGYFLMSPPFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFTVLLFTVACIQTLVLHQQYFHCFFVSGMIRIKTAVIGAYVRKALVITNSARKSSTV 420

Db 361 MQGYFTVLLFVACIQTLVLHQYFHI CFVSGNRIRKTAIVIGAVYRKALVITNSARKSSTV 420  
QY 421 GEIVNLMSVDAQRFMDLATYINNIWSAPLQVILALYLLMLNLGSPSVLAGVAVMLVMPVN 480  
Db 421 GEIVNLMSVDAQRFMDLATYINNIWSAPLQVILALYLLMLNLGSPSVLAGVAVMLVMPVN 480  
QY 481 AVAMKTKTYQVAHMKSKNRIKLMEIINGIKVLKLYAWELAFKDKVLAIRQEBELKVLK 540  
Db 481 AVAMKTKTYQVAHMKSKNRIKLMEIINGIKVLKLYAWELAFKDKVLAIRQEBELKVLK 540  
QY 541 KSAYLSAVCTFTWCTPFLVALCTFAVYVITDENNILLDAQTAFLVSLALENIRFPLNLLP 600  
Db 541 KSAYLSAVCTFTWCTPFLVALCTFAVYVITDENNILLDAQTAFLVSLALENIRFPLNLLP 600  
QY 601 MVISSIVQASVSLKRLIRIFLSHEELEPDSIERPPVKDGGGTNSITVRNATFTWASDDPT 660  
Db 601 MVISSIVQASVSLKRLIRIFLSHEELEPDSIERPPVKDGGGTNSITVRNATFTWASDDPT 660  
QY 661 LINGITPSIEGALVAVVGVGGCKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIQND 720  
Db 661 LINGITPSIEGALVAVVGVGGCKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIQND 720  
QY 721 SURENLLFCQLEEPYRVSIVIOACALLPDLLEILPSGDRTEIGCKGNLSGGQKQVSLAR 780  
Db 721 SURENLLFCQLEEPYRVSIVIOACALLPDLLEILPSGDRTEIGCKGNLSGGQKQVSLAR 780  
QY 781 AVYSNADIYLDPLSAVDHVGKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVIV 840  
Db 781 AVYSNADIYLDPLSAVDHVGKHI FENVIGPKMLKNKTRILVTHSMSYLPQVDVIV 840  
QY 841 MSGGKISEMGYSQELLAROGAFABFLRTYASTEQODAEENGVTGSGPGKEAKQWENG 900  
Db 841 MSGGKISEMGYSQELLAROGAFABFLRTYASTEQODAEENGVTGSGPGKEAKQWENG 900  
QY 901 LVTDSAGLQRLQSSSSYSYSGDI SRHNSSTAELOKAEAKKETWKLMEADKAQOGVKL 960  
Db 901 LVTDSAGLQRLQSSSSYSYSGDI SRHNSSTAELOKAEAKKETWKLMEADKAQOGVKL 960  
QY 961 SVYWDYMKAIGLFISFLSIFLPCNHNVSALSNYLSLWTDPIVNGTOEHTKVLRSVYG 1020  
Db 961 SVYWDYMKAIGLFISFLSIFLPCNHNVSALSNYLSLWTDPIVNGTOEHTKVLRSVYG 1020  
QY 1021 ALGISQGIADVGYSMVAISGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSL 1080  
Db 1021 ALGISQGIADVGYSMVAISGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSL 1080  
QY 1081 DTVDSMIPVIMKFMGSLFNWIGACIVILLATPIAAIIIPPLGLIYFFVQRYFVASSRQL 1140  
Db 1081 DTVDSMIPVIMKFMGSLFNWIGACIVILLATPIAAIIIPPLGLIYFFVQRYFVASSRQL 1140  
QY 1141 KELESVRSRSPVYSHFNETLLGVSVIRAFEEQERFTHQSDLKVDENOKAYPSIVANRWLA 1200  
Db 1141 KELESVRSRSPVYSHFNETLLGVSVIRAFEEQERFTHQSDLKVDENOKAYPSIVANRWLA 1200  
QY 1201 VRLCVCNGLVILFAALFAVISRHSLSAGLVGLSVSYLQVTTYLANWLRMSSEMETNIVA 1260  
Db 1201 VRLCVCNGLVILFAALFAVISRHSLSAGLVGLSVSYLQVTTYLANWLRMSSEMETNIVA 1260  
QY 1261 VERLKEYSETEKAPWQIOETAPPSSWPQVGRVFRNRYCLRYREDLDLVLRHINTYINGG 1320  
Db 1261 VERLKEYSETEKAPWQIOETAPPSSWPQVGRVFRNRYCLRYREDLDLVLRHINTYINGG 1320  
QY 1321 EKVGVGRGAGKSSITLGLFRINSAGEEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380  
Db 1321 EKVGVGRGAGKSSITLGLFRINSAGEEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380  
QY 1381 SGLSRMLNDPFSQYDEEYVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
Db 1381 SGLSRMLNDPFSQYDEEYVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGQRLVCL 1440  
QY 1441 ABALLRKTKILVLDENATAVDLETDLIQSTTRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
Db 1441 ABALLRKTKILVLDENATAVDLETDLIQSTTRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500

QY 1501 DKGEIQEYCAPSDLLQORGLFYSMKADAGL 1530  
Db 1501 DKGEIQEYCAPSDLLQORGLFYSMKADAGL 1530

## RESULT 15

AAR54928  
ID AAR54928 standard; protein; 1531 AA.

XX AAR54928;

DT 25-MAR-2003 (revised)  
DT 14-OCT-1994 (first entry)

XX Multidrug resistance protein.

XX Multidrug resistance protein; MRP; H69AR; cancer cell line; stem cell;  
cardiac muscle; transgenic animal.

XX Homo sapiens.

XX WO9410303-A1.

XX 11-MAY-1994.

XX 27-OCT-1993; 93WO-CA000439.

XX 27-OCT-1992; 92US-00966923.

XX 08-MAR-1993; 93US-00029340.

XX (TOOH ) UNIV QUEBENS KINGSTON.

XX Deeley RG, Cole SPC;

XX WPI; 1994-167460/20.

XX N-PSDB; AAQ65377.

XX Multi-drug resistance gene - encodes protein capable of conferring multi-  
drug resistance on cells, useful in diagnostic and treatment methods.

XX Disclosure; Page 69-74; 101pp; English.

XX The multidrug resistant cancer cell line H69AR (ATCC CRL 11350) was used  
to identify cDNA encoding a novel protein associated with multidrug  
resistance, MRP. MRP may be expressed in e.g. hematopoietic stem cells or  
cardiac muscle, or in transgenic animals, or can be used to raise  
antibodies. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 1531 AA;

Query Match 78.4%; Score 7849; DB 2; Length 1531;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCADGSDPLDNDNNTWNTSNPDKFQNTLVWVPCFYLWACFPFYLYLSRH 60

Db 1 MALRGFCADGSDPLDNDNNTWNTSNPDKFQNTLVWVPCFYLWACFPFYLYLSRH 60

QY 61 DRGYIQMTPLNKTALGFLWVCWADLFYSFWRSGIFLAPVFLVSPITLLGITTLA 120

Db 61 DRGYIQMTPLNKTALGFLWVCWADLFYSFWRSGIFLAPVFLVSPITLLGITTLA 120

QY 121 TFLQLERKGVQSSGIMLTFWLVALCALAILRSKINTALKEDAQVDLFRDITFVYVFS 180

Db 121 TFLQLERKGVQSSGIMLTFWLVALCALAILRSKINTALKEDAQVDLFRDITFVYVFS 180

QY 181 LLLIQVLVSCFSDRSPLESETIHDNPNCPSSASFLSRITFWITGLIVRGVQPLEGSD 240

Db 181 LLLIQVLVSCFSDRSPLESETIHDNPNCPSSASFLSRITFWITGLIVRGVQPLEGSD 240

QY 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKQPKVYVSSKDPAPKSSKVDANESEVAL 300

Db 241 LWSLNKEDTSQVVPVLVKNWKKCAKTRKQPKVYVSSKDPAPKSSKVDANESEVAL 300



Db 241 LWSLNKEDTSEQVVPVVLVKNWKECAKTRKQPVKVYVSSKDPQPKSSKVDANEVEAL 300  
 QY 301 IVKSPKQKNPSLFKVLKTFGYPFLMSFFKAIHDLMMFSGPOIILKLIKFNVDTKAPD 360  
 Db 301 IVKSPKQKNPSLFKVLKTFGYPFLMSFFKAIHDLMMFSGPOIILKLIKFNVDTKAPD 360  
 QY 361 WQGYFYTVLLFVACLTQLVHLQYFHCIFVSGMRKIKTAVIGAVYRKALVITNSARKSSTV 420  
 Db 361 WQGYFYTVLLFVACLTQLVHLQYFHCIFVSGMRKIKTAVIGAVYRKALVITNSARKSSTV 420  
 QY 421 GETVNLMSVDAQRFMDLATYINNIWSAPLOVILALYLLMLNIGPSVLAVVWLMVNVN 480  
 Db 421 GETVNLMSVDAQRFMDLATYINNIWSAPLOVILALYLLMLNIGPSVLAVVWLMVNVN 480  
 QY 481 AVWAMTKTYQVAHMKSKDNRIKLNNEIILNGIKVLYAWELAFKDKVLAIQEBELKVLK 540  
 Db 481 AVWAMTKTYQVAHMKSKDNRIKLNNEIILNGIKVLYAWELAFKDKVLAIQEBELKVLK 540  
 QY 541 KGAYLSAVGTFTWVCTPFLVALCTFAVYVTTIDENNILDAQTAFVSLALFNILRFPNLILP 600  
 Db 541 KGAYLSAVGTFTWVCTPFLVALCTFAVYVTTIDENNILDAQTAFVSLALFNILRFPNLILP 600  
 QY 601 WVISSIVQASVSLKRLRIFLSHHELEPDSIERRPVKDGGGTNSITVRNATFTWASDPPT 660  
 Db 601 WVISSIVQASVSLKRLRIFLSHHELEPDSIERRPVKDGGGTNSITVRNATFTWASDPPT 660  
 QY 661 LINGITFSIPEGALVAVVGQVCGKSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIOND 720  
 Db 661 LINGITFSIPEGALVAVVGQVCGKSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIOND 720  
 QY 721 SURENILFCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKGVNLSGGQKQVSLAR 780  
 Db 721 SURENILFCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEKGVNLSGGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSADVAHVCKHI FENVIGPKMLKNKTRILVTHSMYLPQVDVIV 840  
 Db 781 AVYSNADIYLFDDPLSADVAHVCKHI FENVIGPKMLKNKTRILVTHSMYLPQVDVIV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFELRTYASTEQDQDAEENGVTGVSFGPKAKOMENG 900  
 Db 841 MSGGKISEMGSYQELLARDGAFELRTYASTEQDQDAEENGVTGVSFGPKAKOMENG 900  
 QY 901 LVYDSAGLQRLQSSSSSYSGDISRHNSHSTAELQAKBAKKEETWKLMEADKAOTGVKL 960  
 Db 901 LVYDSAGLQRLQSSSSSYSGDISRHNSHSTAELQAKBAKKEETWKLMEADKAOTGVKL 960  
 QY 961 SVYDYMKAIGLFI SFLIFLPMCNHVSALASNYWLSLWTDPIVNGTQHTKVRLSVYG 1020  
 Db 961 SVYDYMKAIGLFI SFLIFLPMCNHVSALASNYWLSLWTDPIVNGTQHTKVRLSVYG 1020  
 QY 1021 ALGISQGIAGVGYSMASVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSEL 1080  
 Db 1021 ALGISQGIAGVGYSMASVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFKSEL 1080  
 QY 1081 DTVDSMIPEVIMKFMGSLFNIGACIVILLATPIAAIIIPGLIYFPVQRFYVASSRQL 1140  
 Db 1081 DTVDSMIPEVIMKFMGSLFNIGACIVILLATPIAAIIIPGLIYFPVQRFYVASSRQL 1140  
 QY 1141 KRLESVSRSPVYSHENETLLGVSVIRAFEEQERFTHQSDLVKVDENOKAYPSIVANRWLA 1200  
 Db 1141 KRLESVSRSPVYSHENETLLGVSVIRAFEEQERFTHQSDLVKVDENOKAYPSIVANRWLA 1200  
 QY 1201 VRLECVGNICIVLFAALFAVISRHSLSAGLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
 Db 1201 VRLECVGNICIVLFAALFAVISRHSLSAGLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKAPWQIQTAPSSWPQGVFRNRYCLRYREDLDFVLRHINVTINGG 1320  
 Db 1261 VERLKEYSETEKAPWQIQTAPSSWPQGVFRNRYCLRYREDLDFVLRHINVTINGG 1320  
 QY 1321 EKVGI VGRGTAGKSSLTGLFRINESAECEIIIDGINIAKIGLHDLRPFKTIIPQDPVLF 1380  
 Db 1321 EKVGI VGRGTAGKSSLTGLFRINESAECEIIIDGINIAKIGLHDLRPFKTIIPQDPVLF 1380

QY 1381 SGSLRMNLDPPFSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGORQLVCL 1440  
 Db 1381 SGSLRMNLDPPFSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAGGENLSVGORQLVCL 1440  
 QY 1441 ARALLRKTKILVLDATAAVALDETDDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTKILVLDATAAVALDETDDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIOEYCAPSDLLQORGLFYSMADAGLV 1531  
 Db 1501 DKGEIOEYCAPSDLLQORGLFYSMADAGLV 1531

Search completed: March 18, 2005, 11:02:20  
 Job time : 146.718 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:20 ; Search time 45.7349 seconds  
(without alignments)  
4096.085 Million cell updates/sec

Title: US-10-665-283-8  
Perfect score: 10016  
Sequence: 1 MALRGFCSADGSDPLMDWNV.....RSVAVAKPKFSISPDLS 1947

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	7849	78.4	1531	1	DVHUAR
2	4484.5	44.8	1527	2	multidrug resistan
3	3587	35.8	1545	1	canalicular multis
4	3507.5	35.0	1541	1	multidrug resist
5	3289	32.8	1494	2	canalicular multid
6	3225	32.2	1573	2	protein F57C12.4 [
7	3195.5	31.9	1502	2	hypothetical prote
8	2838	28.3	1515	1	multidrug resist
9	2732	27.3	1398	2	cadmium resistance
10	2548	25.4	1478	2	hypothetical prote
11	2384.5	23.8	1623	2	ABC transporter SP
12	2356.5	23.5	1622	2	ABC transporter At
13	2336	23.2	1495	2	glutathione S-conj
14	2312	23.1	1144	2	probable ABC trans
15	2294.5	22.9	1559	1	hypothetical prote
16	2263.5	22.6	1488	2	probable ABC trans
17	2238.5	22.3	1516	2	glutathione-conjug
18	2230	22.3	1539	2	ABC transporter-11
19	2144.5	21.4	1355	2	hypothetical prote
20	2144.5	21.4	1514	2	multi resistance p
21	2140	21.4	1515	2	MRP-like ABC trans
22	2131.5	21.3	1490	2	multi resistance p
23	2129	21.3	1545	2	sulfonylurea recep
24	2123	21.2	1153	2	hypothetical prote
25	2109	21.1	1545	2	sulfonylurea recep
26	2104.5	21.0	1511	2	sulfonylurea recep
27	2098	20.9	1546	2	sulfonylurea recep
28	2069.5	20.7	1389	2	ABC transporter-11
29	2062.5	20.6	1661	2	probable membrane

ALIGNMENTS

RESULT 1

DVHUAR

multidrug resistance protein (cell line H69AR) - human  
N:Alternate names: multidrug resistance-associated protein (MRP)  
C:Species: Homo sapiens (man)  
C>Date: 30-Jun-1993 #sequence\_revision 05-Dec-1998 #text\_change 13-Jan-2001  
C:Accession: A44231; A37495  
R:Coile, S.P.C.; Bhardwaj, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almquist, K.C.;  
Science 258, 1650-1654, 1992  
A:Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer.  
A:Reference number: A44231; MUID:93088080; PMID:1360704  
A:Accession: A44231

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 'MAPTRSGTMSRGIPATPTSPAFRRSSCGLVFTSGPV', 50-1531 <Col>

A:Cross-references: GB:L05628; NID:gl835658

A:Experimental source: small cell lung carcinoma cell line H69AR

A>Note: sequence extracted from NCBI backbone (NCBIP:119851); this sequence has been corrected.  
R:Coile, S.P.C.; Deeley, R.G.  
Science 260, 879, 1993

A:Title: Multidrug resistance-associated protein: sequence correction.

A:Reference number: A37495; MUID:93262415; PMID:8098549

A:Accession: A37495

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-60 <CO2>

A:Cross-references: GB:L05628; NID:gl835658

A>Note: sequence extracted from NCBI backbone (NCBIP:131929)

C:Genetics:

A:Gene: GDB:MRP

A:Cross-references: GDB:136335; OMIM:158343

A:Map position: 16p13.1-16p13.1

C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

C:Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transmembrane

F:661-844/Domain: ATP-binding cassette homology <ABC1>

F:678-685/Region: nucleotide-binding motif A (P-loop)

F:788-792/Region: nucleotide-binding motif B

F:1310-1503/Domain: ATP-binding cassette homology <ABC2>

F:1327-1334/Region: nucleotide-binding motif A (P-loop)

F:1450-1454/Region: nucleotide-binding motif B

Query Match 78.4%; Score 7849; DB 1; Length 1531;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDWNVNTNTSPDFTKCFQNTLVWVPCFYLWACFPFYLSRH 60

Db 1 MALRGFCSADGSDPLMDWNVNTNTSPDFTKCFQNTLVWVPCFYLWACFPFYLSRH 60

QY 61 DRGYQMTPLNKTALGFLMWICWADLFYSFWSRSGIFLAPVFLVSPITLLGITLLA 120

Db 61 DRGYQMTPLNKTALGFLMWICWADLFYSFWSRSGIFLAPVFLVSPITLLGITLLA 120

QY 121 TFLIQLERRKGVOSSGIMLTFWLVALYCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TFLIQLERRKGVOSSGIMLTFWLVALYCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 QY 181 LLLIQLVLSCFSDRSPLFSTIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 DB 181 LLLIQLVLSCFSDRSPLFSTIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSD 240  
 QY 241 LWSLINKEDTSEOVVPLVLRNKKCECAKTRKQPVKVVYSSKDPAPKPESSKVDANEVEAL 300  
 DB 241 LWSLINKEDTSEOVVPLVLRNKKCECAKTRKQPVKVVYSSKDPAPKPESSKVDANEVEAL 300  
 QY 301 IVKSPQKEMNPFLFKVLYKTFPGVFLMSFPFKAIHDLMMFSGPQIILKLIKFNVDTKAPD 360  
 DB 301 IVKSPQKEMNPFLFKVLYKTFPGVFLMSFPFKAIHDLMMFSGPQIILKLIKFNVDTKAPD 360  
 QY 361 WQGYFYTVLLFVTFACLTQLVHGYFHCIFVSGMIRIKTAVIGAYYRKALVITNSARKSSTV 420  
 DB 361 WQGYFYTVLLFVTFACLTQLVHGYFHCIFVSGMIRIKTAVIGAYYRKALVITNSARKSSTV 420  
 QY 421 GETVNLMSVDAQRFMDLATVINNIWSAPLOVILALYLLNLGSPVLGAVVAVMLVMPVN 480  
 DB 421 GETVNLMSVDAQRFMDLATVINNIWSAPLOVILALYLLNLGSPVLGAVVAVMLVMPVN 480  
 QY 481 AVNMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQEEELKVLK 540  
 DB 481 AVNMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIRQEEELKVLK 540  
 QY 541 KSAYLSAVGTFTWCPTFLVACLTFAVYVITDENNILDQAOTAFVSLALFNILRFPNLILP 600  
 DB 541 KSAYLSAVGTFTWCPTFLVACLTFAVYVITDENNILDQAOTAFVSLALFNILRFPNLILP 600  
 QY 601 MWISSIVQASVSLKRLRIFLSHEELPDSIERPVKDGGSITSITVRNATFTWASDDPT 660  
 DB 601 MWISSIVQASVSLKRLRIFLSHEELPDSIERPVKDGGSITSITVRNATFTWASDDPT 660  
 QY 661 LINGITFSEPALVAVGVGCGKSLLSALLAEMDKVEGHVAIRKGSVAYVPQAWIQND 720  
 DB 661 LINGITFSEPALVAVGVGCGKSLLSALLAEMDKVEGHVAIRKGSVAYVPQAWIQND 720  
 QY 721 SURENILFCQLEEPYRVSIVOCALLPDLPLPSGDRTEIGEGKGNVLSGGQKQVSLAR 780  
 DB 721 SURENILFCQLEEPYRVSIVOCALLPDLPLPSGDRTEIGEGKGNVLSGGQKQVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSADVAHVGHKHFENVIGPKGMLKNKTRILVTHSMYLPQVDVLIIV 840  
 DB 781 AVYSNADIYLFDDPLSADVAHVGHKHFENVIGPKGMLKNKTRILVTHSMYLPQVDVLIIV 840  
 QY 841 MSGGKISEMGSYQELLARDGAPAEFLRTVASTEQDQAEENGVTGVSGFGKEAKQWENG 900  
 DB 841 MSGGKISEMGSYQELLARDGAPAEFLRTVASTEQDQAEENGVTGVSGFGKEAKQWENG 900  
 QY 901 LVYDSAGKQLORQLSSSSSYSGDISRHNSSTAELQAKAEKEETWKLMEADKAQGTQVKL 960  
 DB 901 LVYDSAGKQLORQLSSSSSYSGDISRHNSSTAELQAKAEKEETWKLMEADKAQGTQVKL 960  
 QY 961 SVYDYMKAIGLIFSLFIFLPMCNHVSALASNYLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 DB 961 SVYDYMKAIGLIFSLFIFLPMCNHVSALASNYLSLWTDPIVNGTQEHTKVRLSVYG 1020  
 QY 1021 ALGISQGTAVFGYSMAVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRNPSKEL 1080  
 DB 1021 ALGISQGTAVFGYSMAVIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVRNPSKEL 1080  
 QY 1081 DTVDMSIEVIMKFMGSLFNIVIGACIVILLATPTAAIIIPPLGLIYFPVQRYFVASSROL 1140  
 DB 1081 DTVDMSIEVIMKFMGSLFNIVIGACIVILLATPTAAIIIPPLGLIYFPVQRYFVASSROL 1140  
 QY 1141 KRLESVSRSPYSHNETLLGVSVIRAPAEQERFTHQSDLKVDENQKAYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPYSHNETLLGVSVIRAPAEQERFTHQSDLKVDENQKAYPSIVANRWLA 1200

QY 1201 VRLECVGNCIVLPAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMENIVA 1260  
 DB 1201 VRLECVGNCIVLPAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVMSSEMENIVA 1260  
 QY 1261 VERLKEYSTEKAPWQIOETAPPSSWPQVGRVERPNYCLRYREDLDVFLRHINVTINGG 1320  
 DB 1261 VERLKEYSTEKAPWQIOETAPPSSWPQVGRVERPNYCLRYREDLDVFLRHINVTINGG 1320  
 QY 1321 EKVGIVGRTGACKSSITLGLFRINSAEGEIIGDINIAKIGLHDLRFKTIIPQDPVILF 1380  
 DB 1321 EKVGIVGRTGACKSSITLGLFRINSAEGEIIGDINIAKIGLHDLRFKTIIPQDPVILF 1380  
 QY 1381 SCSLRNLDLPFQYDEEVTSLAHLKDFVSALPDKLDHCEAGGENLSVGQRLVCL 1440  
 DB 1381 SCSLRNLDLPFQYDEEVTSLAHLKDFVSALPDKLDHCEAGGENLSVGQRLVCL 1440  
 QY 1441 ABALLARKTKILVLDATAVLETDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
 DB 1441 ABALLARKTKILVLDATAVLETDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500  
 QY 1501 DRGETQYEGAPSDLLQORGLFYSMADAGLV 1531  
 DB 1501 DRGETQYEGAPSDLLQORGLFYSMADAGLV 1531  
 RESULT 2  
 JB0336  
 canalicular multispecific organic anion transporter - human  
 C:Species: Homo sapiens (man)  
 C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 09-Jul-2004  
 C:Accession: JB0336  
 R:Uchiumi, T.; Hinochita, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Toh, S.; Furukawa, M.;  
 Biochem. Biophys. Res. Commun. 252, 103-110, 1998  
 A:Title: Isolation of a novel human canalicular multispecific organic anion transporter,  
 t.  
 A:Reference number: JB0336; MUID:99032812; PMID:9813153  
 A:Accession: JB0336  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1527 <UCH>  
 A:Cross-references: UNIPROT:O15438; GB:AF089352  
 A:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
 C:Keywords: ATP  
 F:1306-1499/Domain: ATP-binding cassette homology <ABC2>  
 Query Match 44.8%; Score 4484.5; DB 2; Length 1527;  
 Best Local Similarity 56.6%; Pred. No. 3.2e-293;  
 Matches 871; Conservative 261; Mismatches 372; Indels 35; Gaps 9;  
 QY 8 SADGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFVLWACPPFYFLYLSRHDGRGYIQM 67  
 DB 7 SGEIGSKFWSNLSVHTENPDLTFCFQNSLLAWVPCYLVWVALPCYLLYLRRHRCRGIIL 66  
 QY 68 TPLNKTATGLFLWIVCWADLFYSFWERSRGIFLAPVFLVSPTLGTTLLATLFLIQLE 127  
 DB 67 SHLSKLMVLGVLLWCVSADLFYSFHLVHGRAPAPVFFVTPLVGVVTMLAFLLIQYE 126  
 QY 128 RRKGVQSSGIMLTFWLVALYCALAILRSKIMTALKEDAQVDLFRDITFYVYFSLILQLV 187  
 DB 127 RLQGVQSSGIMLTFWLVALYCALAILRSKIMTALKEDAQVDLFRDITFYVYFSLILQLV 186  
 QY 188 LSCFSDRSPLFSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLEGSDLWSLNKE 247  
 DB 187 LACFRKPPFPFSAKNVDNPNYPETSAGLSRLFWFTKMAIYGYRHPLEKDLWSLKEE 246  
 QY 248 DTSEQVVPVLVKNKKCECAKTRKQPVKVVYSSKDPAPKPESSKVDANEVEALIVKSPQK 307  
 DB 247 DRSQMVVQQLLEARWKOEKQOTARH-----KASAPGK---NASGEDEVILGARPRP 294  
 QY 308 EWNPSLFLKLYKTFPGVFLMSFPFKAIHDLMMFSGPQIILKLIKFNVDTKAPDQGVFYT 367  
 DB 295 R-NPSFLKALLAIFGSSFLISACFKLIQDLISPINQLLSILIRFISNPNGPSWGWFLVA 353

```

Db      1428 VGORQVCLARALLRSRILVLEDEATAADLETDNLIQATIRQFTDTCTVLTIAHRLNTI 1487
QY      1492 MDYTRVILVDKGBIQBYGAPSDLLQOQGLFYFMAKDAGL 1530
Db      1488 MDYTRVILVDKGVVAEFDSPANLIARGIIFYGNARDAGL 1526

RESULT 3
S71841
multidrug resistance protein, canalicular - human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C;Accession: S71841; S71840
R;Koening, J.; Keppeler, D.
submitted to the EMBL Data Library, August 1996
A;Reference number: S71841
A;Accession: S71841
A;Molecule type: mRNA
A;Residues: 1-1545 <KOE>
A;Cross-references: EMBL:X96395; NID:g1507819; PIDN:CAA65259.1; PID:g1514568
R;Buechler, M.; Koening, J.; Brom, M.; Kartenbeck, J.; Spring, H.; Horie, T.; Keppeler, D.
J. Biol. Chem. 271, 15091-15098, 1996
A;Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance protein
A;Reference number: S71839; MUID:96279006; PMID:8662992
A;Accession: S71840
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1415-1429,'VP',1432-1455,'E',1457-1545 <BUE>
A;Cross-references: EMBL:X96395
C;Genetics:
A;Gene: GDB:ABCC2; CMOAT; ABC; MRP2; CMRP; DJ5
A;Cross-references: GDB:6089489; OMIM:601107
A;Map position: 10q24-10q24
C;Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette
C;Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein
F;654-837/Domain: ATP-binding cassette homology <ABC1>
F;671-678/Region: nucleotide-binding motif A {P-loop}
F;1317-1510/Domain: ATP-binding cassette homology <ABC2>
F;1334-1341/Region: nucleotide-binding motif A {P-loop}

Query Match      35.8%; Score 3587; DB 1; Length 1545;
Best Local Similarity 47.5%; Pred.No. 9.7e-233;
Matches 746; Conservative 286; Mismatches 461; Indels 78; Gaps 20;

QY      3 LRGFCSADGSDPLMDWNVTWNTS-----NPDFTKCFQNTVLVWVPCFYLWACPFVFL-- 55
Db      2 LKFCN-----STFWNSSFLDSPADLPCLCFEQTVLVWIPGLFLLAPWQLLHV 51
QY      56 YLSHRDGYIQTPLNKTKTALGFLIWIWCVADLFVSFWERSGIFLAPVFLVSPDLLGI 115
Db      52 YKSRTRKSRSTTKLYLAK-QVVFVGLL-ILAAIELALVLTEDSGQATVPAVRYTNPSLY-L 108
QY      116 TTLLATFLIQLERRKGVSQSGIMLT-FWLVALVALAILRSKIMTALKEDAQVDLFRDIT 174
Db      109 GTWLLVLLIIQYSQWCQVKNSWFLSLFWILSIICGTFFQFQFTLRTLLOGD-NSNLAYSCL 167
QY      175 FYYVFSLLLQLVLSCFSDRSLPSETIHDPNCPPESSASFLSRITFWITGLIVRGYRQ 234
Db      168 FFISYGFQIILIFSAFSEN-----NBSSNNPSSIASFLSSITVSYWYDSIILKGYKR 219
QY      235 PLGSDLSLWLNKEDTSEQVVPVLVKNWKKCEAKTRK-----QPVKVYVSSKD 281
Db      220 PLETEDVWEDEEMKTKTLVSKFETHMKRLQKARALQRRQSKSQONSAGALPLGLNKN 279
QY      282 PAQPKSSKVDANEEVEALIVKSPQKEWNPSS--LFKVLKYTKFGPYFLMSFFFKAIHDLMM 339
Db      280 QSQSDALVL---EDVEKKKKSGTKQDVPKSWLMKALFKTFYVWLLKSFLLKLVNDIFT 336
QY      340 FSGPQILKLLIKFVNDTKAPDWOGYFVTVLLFVTAQLTVLHVOYFHCIFVSGMRKLTAV 399
Db      337 FVSPOLLKLLISASDRDTLWITGYLCAILLFTAAITQSPCLQCYFOLCFKGLGVKVRTAI 396

```



```

Db      344 GFVKSNSYVWFYICAILMFVAVTLIQSFCLQSYQHCFVLGVCVRITVMSIYKALTL 403
Qy      411 TNSARKSVTVGINVMSVDAQFVMDLATYINNIWSAPLOVTLALYLWNLGSPVLGAV 470
Db      404 SNLARKQYITIGETVNLMSVDSQKMDATYMQVWSSVQITLSIFFLRELGPISILAV 463
Qy      471 AVMLVMPVNAVMAKTKTYQVAHMKSKDNRIKLAMNEILNGIKVLKYLAWELAFKDKVLA 530
Db      464 GVMVLLIPVNGVLATKIRNIQVNMKNKDKRLKINNEILSGIKILKYFAWEPSPQVQV 523
Qy      531 IQBELKVLKKSAYLSAVGTFWVCTPPLVALCTFAVYVITIDENNILDQATFVSLALFN 590
Db      524 IRKELKNLRFQQLQSLIFLQITPILVSVVTFVSVYVLDVDSANVNAEKAFITILFN 583
Qy      591 ILRFPNLPMVSISSVQASVSLKRLRIFLSHEELEPDSIBRRPVKGGGTWSITVRNAT 650
Db      584 ILRFPNMLPMTSSILQASVSDRLERYLGGDDLDTSAIRVSNFD----KAVFSEAS 639
Qy      651 FTWARSDDPTLNGITPISPEGALVAVGVQVGGCKSSLSALLAEMDKVEGHVAIRGSVAY 710
Db      640 FTWDPLEATIQVNDLIKPGQLVAVGVTVGSKSSLSVSMGLAMENVHGHTIQGSTAY 699
Qy      711 VPQAWIQWDSREHILFCQLEBPYRSVIOACALLPDLEILPGSDRTEIGEGVNLG 770
Db      700 VPQSWIQNGTIKDNILFGSEYNEKYOQVLKACALLPDLEILPGDMAEIGEGINLSG 759
Qy      771 GOKQVSLARAVYSNADYFLDDPLSADVAHVGHIFENVIGPGLMKLNKTRILYTHSMS 830
Db      760 GOKQVSLARAYQADYIYLDPLSADVAHVGHIFNKVGPNGLLAGKTRIFVTHGH 819
Qy      831 YLPQVDVIVMSGGKISEMGYSQELLARDGAFABFLRTYAS-TEQEQDAEENGVTGVSQP 889
Db      820 FLQVDEIVVLGKGTILEKSYRDLLDKGVFARNWKTFMKHSQGPGEATVNN----- 872
Qy      890 GKEAKQMGMLVT-----DSAG-----KOLQQLSSSSSYSGDISRHHSNTEALQ-- 935
Db      873 DSEAEEDDDGLIPTMEIEPEDAASLAMRRENSLRRTLSRSSSSRRRGKSLNKLKINV 932
Qy      936 ---KAEAKKEETWKLMEADKATQGVKLSVYVDMYKATGLFSLFPLFCMNHVSALAS 992
Db      933 NVLKEKEVEGOKLIKKEFEVTKGVKFSYIKYIQVAGVWMBILFIILFYGLNNVAFGS 992
Qy      993 NYWLSLWT--DDPIVNGT---QBHTKVLRSVYVYALGISTQIAVFGYSMAVSIIGILASRL 1048
Db      993 NLWLSAWTSDSNLNGTNNSSSHRDMRIGVFGALGLAQICLLISTLSIYACRNASKAL 1052
Qy      1049 HVDLLHSILRSPMSFPTPSGNLVNRFSEKELDTVDSDMIPKVMFMSGLFNVIACIVI 1108
Db      1053 HQQLLTNLRAPMRFPDTPPTGRIVNRFSGDITVDLLPQTLRSWMMCFFGIAGTLVMI 1112
Qy      1109 LLAATPAAIIIPGLGIYFFQVRYVASSRQLKLESYRSRSPVYSHFNETLLGVSVIRAF 1168
Db      1113 CMATPVFAIIIPLSILYISQVYVYATSRQRRLDSVTKSPISYHFSFTVGLPIIRAF 1172
Qy      1169 EQERFIHOSDLKVDENOKAYPSIVANRWLAVRLECVGNCIVLPAALFAVTSRHSLSAG 1228
Db      1173 EQHQFLAWNEKQIDINQKCVSWITSNWLAIKRLVGNLVVFCALLVYRKLTCGD 1232
Qy      1229 LVGLSVSVSQVTVYVNLVNLVMSMETNIVAVERLKEYSETEKAPQIOETAPSSWP 1288
Db      1233 VVGFLSNALNITOTLNLVLRMTSEATNIVAVERTSEYINVENEAPW-VTDKRPADWP 1291
Qy      1289 QVGRVEFRNYCLRYREDLDFVLRHINVTINGEKYIGVTRTGAGKSSLTGLFRINESAE 1348
Db      1292 RHGEIFQNNYQVRYPELDELVLKGTICNTKSGEKVGVWGRGTGAGKSSLTNCLFRILESAG 1351
Qy      1349 GRIIIDGINIAIGHDLRFKTIIPQDPVLFSGSLRMLNLPFFSOYSDSEVWTSLELAHL 1408
Db      1352 GOIIGDIDVAGIGHDLRERITIIIPQDPILFSGSLRMLNLPFNKYSDEEVRWALELAHL 1411
Qy      1409 KDFVSLPDKLDECAEGENISVGORQLVCLARALLRKTILVLDATAAVDLTDDLI 1468

```

```

Db      1412 RSFVSGLOGLLSEVTEGGDNLSIQORQLCLGRVLRKSKILVLDATAAVDLTDSLI 1471
Qy      1469 OSTRTQPEDCTVLTIAHRLNTIMDYRVIVLDKGEIOBYGAPSDLLQORGLFYMAKDA 1528
Db      1472 QTTIRKESQCTVITIAHRLTIMDSKIMVLDNGKIVEYGSPELLSNRGSFYLMAKEA 1531
Qy      1529 GL 1530
Db      1532 GI 1533

RESULT 5
E89447
protein F57C12.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: E89447
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_e
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A:Accession: E89447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1494 <STO>
A:Cross-references: UNIPROT:Q20943; GB:chr_X; PIDN:AAA83299.1; PID:g1118071; GSPDB:GN00C
C:Gene: F57C12.4
C:Map position: X
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match      32.8%; Score 3289; DB 2; Length 1494;
Best Local Similarity 47.6%; Pred. No. 1.1e-212;
Matches 696; Conservative 239; Mismatches 435; Indels 92; Gaps 18;

Qy      119 LATLIQIERRRGVSSGIMLTFWLVALVLCALAILRSKIMTALKEDAQVDLFRDITFYVY 178
Db      76 LALILTVACKNGIITSGVITLYWLVVCGIPEFRFVLSGFIYNEYALEGIRATLYIITA 135
Qy      179 FSLLIQILVCLSPDSRPLSETHIDPNCPDESSASFLSRITFWITGLIVRGYQPLRG 238
Db      136 FTFSALELFLCCFAD---VPDMYKSESSCEPTASFINRLTFQWFTGLAYLGNKKSLEN 192
Qy      239 SDSLNLKEDSEQVVPVLVKNWKKCAKTRQPKVYVYSSKDP--AOPKSSSKVDANEE 296
Db      193 EDLMDLNEIDKAENLIPSMQNLKPRIDEYH-QNIK-----KPSAALPKN----- 237
Qy      297 VEALIVKSPQKEWNPSPFLVLYKTFPGPYFLMSFFPKAIHDLMMFSGPQILKLLIKFVNDT 356
Db      238 -----HPSFVIPFKTYKTYLLAGFFYKLCFCDMLQFLAPQLLKQLIGFIEDK 284
Qy      357 KAPDHQGYFYTVLLFPVTAQLOTVLVHQFHCIFVSGMIRIKTAVIGAVYRKALVITNSARK 416
Db      285 NOPWIGCSIVGIMPFSSFLQSMFLHQYVYHSMFRIGMHVRSVLTSAVYSKALNLSNEARK 344
Qy      417 SSTVGEIVNLMSVDAQRFMDLATYINMTWSAPLOVTLALYLWNLGSPVLGAVVMLM 476
Db      345 KGTIGAVNLMSVDIQIKYQDMAPTLMFWASAPLQIFLSYIFLWKLFGVAALAGLVLLILA 404
Qy      477 VPVNAVMAKTKTYQVAHMKSKDNRIKLAMNEILNGIKVLKYLAWELAFKDKVLAIRQEBL 536
Db      405 LPVNGLIAIQMKCQTEQMKLKDEBRIKWSSEILNGMKVLYSWERSMENWVLIKIRREL 464
Qy      537 KVLKKSAYLSAVGTFWVCTPPLVALCTFAVYVITID-ENNILDAQATFVSLALFNLRPP 595
Db      465 HILKLSYFMAAIVFSWICAPFLASVISFVYVYLDPENNVLTPEITFVALSFLDIRMP 524
Qy      596 LNLPMVSISSVQASVSLKRLRIFLSHEELEPDSITERPVKGGGTNSITVRNATFTWAR 655
Db      525 LAMVAVMYGEAVQCSVSNTRLKBFPAEEMSPQT-----SISHGETSDSAIEVNGFLFSWSS 580

```







887 -RPRSDAAPVKGS-----TSEAQMPSLDDVEVTGLTAGEDSGVQYGRVKSAATYLSVLA 940  
 970 IGLFISFLSLFMCNHSALSNYLSLWTDPTDIPVNGTQEHKVLRLSVYALGISQGTIA 1029  
 941 VGTPLCTYTLFLFCQVAFSCOGYLSLWADDPVVDGQKMSALRGSIFGLGCLQALG 1000  
 1030 VFGYSMAVSIGGILASRCLHVDLLHSLRSPMSFFERTPSGNLVNRFSEKELDTVDWSIPE 1089  
 1001 LFSMAAVALFGGARASCLLFRSLMDVARSPICGFERTPVGNLLNRFSEKELDTVDWDIDP 1060  
 1090 VIKMFNGSLPNVIGACIVILLATPIAIIPIPLGLIYFFVQRFYVASSRQLKRLSVRS 1149  
 1061 KMTLLTYAFLGLEVLAVSMATPLAIVAILPLMLLYAGFQSLYVATCCQLRLESASYS 1120  
 1150 PVSHEFNETLLGVSVIRAFEEQRETHQSDLKVDENQKAYPSIVANRWLAVRLCEVGN 1209  
 1121 SVCSHLAETFGQSVVRAFOAQGFPTAQDHADMENQRI SFPLRVADRVLAALELLGNG 1180  
 1210 IVLFAALFAVISRHSLSAGLVLSVSYSLQVTTLYLNLVMSSEMETNIVAVERLKEYSE 1269  
 1181 LVFVAATCAVLSKAHLSAGLGSVSAALQVITQLQWVRSWTDLENSMVAVERVQDYH 1240  
 1270 TEKEAPWQIQTAPPSVQVGRVPRNCLYREDLDLFLVLRHINVTINGSGKVGIVGRT 1329  
 1241 TPKEAPWRLPSSAAQPLWPCGGQIEPRDFGLRHPPELPMAVQGVSLKIHAQKVGIVGRT 1300  
 1330 GAGKSLTLGLPRINESARGEIIGDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMNLD 1389  
 1301 GAGKSLTLGLPRINESARGEIIGDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMNLD 1360  
 1390 PFSQVSDERWTSLELAHLKDFVSLAPDKLDHECAGBGENLSVGORQLVCLARALLRKT 1449  
 1361 LLQENTDEGIAALETVOLKAFVTSPLPQOLQVECSQGDLSVGQKQLLCLARALLRKTQ 1420  
 1450 IIVLDEATAVLETDLQSTIRTPEDCTVLTAAHRLNTMDYTRVILDKGIOEVG 1509  
 1421 IIVLDEATAVDPGTEIQOALERWFAOCTVLLIAHRLRSYMNCAVLVMDGQVASEG 1480  
 1510 APSDLLQKGLFYSMKAOAGL 1530  
 1481 SPAQLLAQKGLFYLAQESGL 1501

RESULT 8  
 S51863  
 N/Alternate names: protein YCF1 - yeast (Saccharomyces cerevisiae)  
 C/Species: Saccharomyces cerevisiae  
 C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C/Accession: S51863; A55352; S50233  
 R/Oliver, K.; Harris, D.  
 submitted to the EMBL Data Library, February 1995  
 A/Reference number: S51863  
 A/Accession: S51863  
 A/Molecule type: DNA  
 A/Residues: 1-1515 <OI>  
 A/Cross-references: UNIPROT:P39109; EMBL:Z48179; NID:G665657; PIDN:CAA88217.1; PID:G6656  
 R/Szcyzka, M.S.; Wemmie, J.A.; Moye-Rowley, W.S.; Thiele, D.J.  
 J. Biol. Chem. 269, 22853-22857, 1994  
 A/Title: A yeast metal resistance protein similar to human cyctic fibrosis transmembrane  
 A/Reference number: A55352; MUID:94357936; PMID:7521334  
 A/Accession: A55352  
 A/Status: nucleic acid sequence not shown  
 A/Molecule type: DNA  
 A/Residues: 1-679, 'R', '681-1515 <SZC>  
 A/Cross-references: GB:L35237; NID:G556464; PIDN:AAA50353.1; PID:G556465  
 C/Genetics:  
 A/Gene: SGD:YCF1; MIPS:YDR135C  
 A/Cross-references: SGD:S0002542; MIPS:YDR135C  
 A/Map position: 4R  
 C/Function:  
 A/Description: required for cadmium resistance  
 C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

C/Keywords: ATP; nucleotide binding; P-loop; transmembrane protein; yeast vacuole  
 F;287-308/Domain: transmembrane #status predicted <TM1>  
 F;345-366/Domain: transmembrane #status predicted <TM2>  
 F;421-442/Domain: transmembrane #status predicted <TM3>  
 F;446-467/Domain: transmembrane #status predicted <TM4>  
 F;534-555/Domain: transmembrane #status predicted <TM5>  
 F;558-580/Domain: transmembrane #status predicted <TM6>  
 F;546-829/Domain: ATP-binding cassette homology <ABC1>  
 F;563-670/Region: nucleotide-binding motif A (P-loop)  
 F;951-972/Domain: transmembrane #status predicted <TM7>  
 F;995-1016/Domain: transmembrane #status predicted <TM8>  
 F;1068-1088/Domain: transmembrane #status predicted <TM9>  
 F;1092-1113/Domain: transmembrane #status predicted <TM10>  
 F;1179-1200/Domain: transmembrane #status predicted <TM11>  
 F;1208-1229/Domain: transmembrane #status predicted <TM12>  
 F;1289-1483/Domain: ATP-binding cassette homology <ABC2>  
 F;1306-1313/Region: nucleotide-binding motif A (P-loop)

Query Match 28.3%; Score 2838; DB 1; Length 1515;  
 Best Local Similarity 40.3%; Pred. No. 2.8e-182;  
 Matches 632; Conservative 295; Mismatches 495; Indels 146; Gaps 27;  
 QY 28 DFTKQFQNTLVWVPCFYLWACFPFYF-----LYLSRHD-----RGVIQMTPLNKTKT 75  
 DB 27 DFTQCFIDGVIL-----NLSALFMTFGRLVNLCKKHSIKYRWNLIIVSRVALVLL 81  
 QY 76 ALGFLWLVWVWADLPYSFWERSRGIFLAPVFLVSTLIGTTLTFLIQLERRKGVSS 135  
 DB 82 EIAP-----VSLASLNISKEAEN---FTIVSQYASTMLSLFVALAHWIEYDR--SVVAN 132  
 QY 136 GIMLTFWLVVALCALAILRSKIMTALKEDAQVLDLFRDITFVYFES-----LLLIOLVLS 190  
 DB 133 TVLFPYLFETFGNPAKILNIRHYEG-----IWSGGTGFILTLFQVITC 180  
 QY 191 FS-----DRSPL-----FSETHDNPCESSASFLSRITFWMITGLIVRGYRQPLEG 238  
 DB 181 ASILLEALPKPLMPHQHITLRRKPNPYDSANIFSRITFSWMSGLMKTGYEKLVE 240  
 QY 239 SDLSLNKEDTSEQVVPVLVGNWKECAKTRKQPVVYSSKDPAPKESKVDANEVE 298  
 DB 241 ADLYKLPRNFSSELSQLEKNWENEL-----KQKS----- 271  
 QY 299 ALIVKSPQKWNPSLKVLYKTGPFYFLMSPFFKAIHDLMMFSGPQILKLIKFNVD-- 355  
 DB 272 -----NPSLSWAICTFGSKMLAAFAKHAHDVLAFTQPOLRLIKLFTVDYNS 320  
 QY 356 -----TKAPDMQGYFVTVLLFVTAQLTLVHLQYFHCIFVSGMRIKTA 398  
 DB 321 ERDDHSSLOGFENNHPQKLPVIRGFLIAFAMFLVGFQTQSVLHQYFLNVENTGMVYKSA 380  
 QY 399 VIGAVYRKALVITNSARKSSTVGEIVNLMSVDAQRFMDLATYINMINSAPLOVILALYL 458  
 DB 381 LTALIYQKSLVLSNEASGLSSTGDIVNLMSVDVQKLOTLQWLNLSWGPQIILCYSL 440  
 QY 459 WNLGSPVLGAVAVVLMVPMVNAWMKTKTYQVAHMKSKDNRIKLNELINGKVLKLY 518  
 DB 441 YKLLGNSMWGVIIILVIMPLNSFLMRKQKQKSKYKDERTRVISEILNLIKSLKLY 500  
 QY 519 AWELAFDKVLAIR-QBELKVLKSAVLSAVGTFTWCTPFLVALCTFAVVTVDENNIL 577  
 DB 501 AWEXPYREKLEFEVNNKELNLTKLGCYMAVTSFQFNIVPFLVSCCTFAVAV-YTEDEL 559  
 QY 578 DAQTAFLVALFNILFRPLNLPVMSIVQASVSLARLRIFFLSHEELEPDSIRRRPKVD 637  
 DB 560 TTDLVFPALTLNLLSFLPMIIPVNLNSFEASVIGRLFTFFNEELQPDSDVQLRFPVK 619  
 QY 638 GGGTNSITV-RNATFTWARS--PPTLNGITFISPEGALVAVVGQVCGKSSLSALLAE 694  
 DB 620 NIGDVAINIGDATTFLWKPVEYKVALKNINFOAKGNLCTIVGKVGSGKALLSCMLGD 679  
 QY 695 MDKVEGHVAIKGVAVYVPOQAMQNDLSRENILFGCQLEBPYRSVTSVQACALLPDLILP 754  
 DB 680 LFRVKGATVHGVSVVYVQVPMNGTIVKENILFGHRYDAEFYEKTIKACALITDLAILM 739



755 SGRTRIGKGNLGGKQKORSLARAVYNSADYILFDDPLSAVDHVGKHFENYVGP 814  
 740 DGDGKTLVGEKIGSLGGKARSLARAVYNSADYILFDDPLSAVDHVGKHFENYVGP 799  
 815 GMLKNTKRLVTHSMGYLPOVDVIVVMSGGKISEMGYQELIA-RDGAFAEFILTYASTE 873  
 800 GLLHTTKVLTATNKVSAISADISALLDNGEITQOCTYDEITKADSPMLKLNLYGK-- 857  
 874 QEQDAENGVTGSGPGKAKQWENGMLVTDGSAKQOLQRLSSSSSYSGDISRHNSAE 933  
 858 -----KNGKSNFEGSSSVRESSIPV-EGELEOLQKLNLDLDFGNSDAISLRASDAT 911  
 934 LQKAEAKCBETWKLMEADKAQOQGVKSLVYDYMKA-----ICLFISFLSIFLFCMCHVS 988  
 912 LGSIDFGDDB--NIAKREHQGVKNWYLYEYAKACPKSCVCFILFIVISNLF----- 964  
 989 ALASNYWLSLWTDPIVNGTQHTKVLRLSVYMGALGISQGIAGVGYSMVSI--GGILASRC 1047  
 965 SVMGNVWLKHWSEVNSRYGNSPNAARYLAIFYALGIGSALATLIQITVLVWFCTIHASKY 1024  
 1048 LHVLDLHLSILRSPMSFERTPSGNLVNRFSEKELDTVDSMIPEVIMFGSLFNVIACIV 1107  
 1025 LNLMTNSVLRAPMTFETTPGRILNRFSDNYKVDALLGRFTSQFFVNAVKVPTTIV 1084  
 1108 ILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKLESVSRSPVSHFNETLLGVSVIRA 1167  
 1085 ICATTWQFIPIILSVFYIYQOYVLRYSRLRLDSITRSPYSHFOETLGLLATVRG 1144  
 1168 FEOERFIHQSLDKVDENOKAYPSIVANRWLAVRLEVCVNCICVLPAAFLAV--TSRHSL 1225  
 1145 YSQKRFSSHINOCDINNNMSAFYPSINANRWLAIRLELIGSIIILGAATLSVFLKQGT 1204  
 1226 SAGLVGLSYSLQVTTYNLWLRMSMETNIVAVERLKEYSETEKEAPMOIQETAPPS 1285  
 1205 TAGVGLSLYSLAQITQTLNWIIVMTVEVETIVSVVERKEVADLSEAPLVEGHRPPK 1264  
 1286 SMPQVGRFRNYCLRYREDLDFVLRHINVTINGGKVGIVGRGTGAGKSSLTGLFRINE 1345  
 1265 EMPQGDIFKNVYSTRYPELDVLVLAHNIHKPEKVGIVGRGTGAGKSSLTGLFRINE 1324  
 1346 SAGEIIGIINIAKGLHDLRPFKTIIPQDPVLFGSGSLRMNLDPPSQYSDSEVWTSLEL 1405  
 1325 ASEGNIVIDNIAINEIGLYDLRHLKSIIPQDSQVFEPTVRENIDPINQYTDIAIRALEL 1384  
 1406 AHLKDFVSALP-DKLDEHCEAGENLSVGOROLCLARALLKTKTLVLDEATAVDLET 1464  
 1385 SHLKEHVLMSNDGLDAQTEGGNLSVGORQLCLARAWLVPFSLTLVLDEATAVDVET 1444  
 1465 DDLIQSTIRTPQEDCTVLTIAHRLNTIMDYTRVILDKGEIOBYGAPSDLL-QORGLFYS 1523  
 1445 DKVQVETITAFKORTILIAHRLNTIMDSRIIVLDNKGVAEPDPSGQLLSDNKSLEFYS 1504  
 1524 MAKADGLV 1531  
 1505 LCWEAGLV 1512

RESULT 9

T20434  
 hypothetical protein E03G2.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T20434  
 R:McMurray, A.  
 submitted to the EMBL Data Library, November 1995  
 A:Reference number: Z19275  
 A:Accession: T20434  
 A:Status: preliminary;  
 A:Molecule type: DNA  
 A:Residues: 1-1398 <WIL>  
 A:Cross-references: UNIPROT:Q19048; EMBL:Z68113; PIDN:CAA92148.1; GSPDB:GN00028; CESP:EO  
 A:Experimental source: clone E03G2

C:Genetics:  
 A:Gene: CESP:E03G2.2  
 A:Map position: X  
 A:Introns: 73/3; 113/3; 176/3; 216/1; 258/2; 384/3; 470/3; 549/1; 644/3; 757/3; 764/2;  
 C:Superfamily: human multidrug resistance protein cMOAT; ATP-binding cassette homology  
 Query Match 27.3%; Score 2732; DB 2; Length 1398;  
 Best Local Similarity 41.1%; Pred. No. 3.4e-175;  
 Matches 591; Conservative 287; Mismatches 472; Indels 88; Gaps 19;  
 QY 115 ITTLATLFIQLERKGVQSSGIMLTFWLVALVCAIALILRSKIMTALKED--AQVDLPD 172  
 DB 27 LALLIATNEV---RRAGIHSSGLPFCIMWLFVAAPVEFYQMTTGSQBELVARIDFFY 83  
 QY 173 IFYVYVFSLLIQLVSCFSDRSPFSETHDPN-----PCPESSASFLSRITFWMITGL 227  
 DB 84 VAYLYFFLVAEFVLFVDFDFPM-----PRGYNLKCPEENANFISRQLLLWFTOI 136  
 QY 228 IVRGYRQPLEGSDLSLNKEDTSEQVVPVLVKNWKECAKTRKQPKV-----VYSS 279  
 DB 137 ISLGYERTLVADDVPEDMSQMDQEVYKARWTEMLKQTEKAREKQVCLDDKREARTGSE 196  
 QY 280 KDPAPQKSSKVDAN--BEVEALIVKSPQKWNPSLFKVLKTFGPFYFLMSFFFAIHLDM 338  
 DB 197 KAPLLGTFTNNYGAVNLDDKORVIVQ-----PSVIVTLWQIMKWBEILGGSEIKFLSLL 249  
 QY 339 MFSGPOILKLLIKFVNDTKAPDQGVFVTVLLFVTAQLQTLVHOFHICFVSGMRIKTA 398  
 DB 250 QPANTFLNYLLIFETPNAPLINGIGLAVGLFLAGQIKSLFMNTYFIAMTRVGAKIQTM 309  
 QY 399 VIGAYRKALVITNSARKSSTVGEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYL 458  
 DB 310 LSCAVYKSLLSNTARRERTVGMVNILSIDVDRFRMITPOIQVWSSPFOIICMWLL 369  
 QY 459 WNLGSPVLAVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAV 518  
 DB 370 SQTIGVAVWAGIVVMSIPINICVSMITKQMLKMYKDBRIELINEVLNGIKVVKLS 429  
 QY 519 AWELAFKDKVLAIROELKLVKKSAYLSAVGTFTVCTPFLVALCTFAVVTIDENNILD 578  
 DB 430 AWETAMEETIERVRDEKELKXIKQSALLKTADCLNAGAPVVALSSFTVFLIDPKNVLT 489  
 QY 579 AQTAFVSLALFNILFPNLPWISSIVQASVSLKRLRIFLSHHEELPDSIERPVDKG 638  
 DB 490 PNIAFVSLSLFLLRGLPMAAELVAQTQVLSVSNKVRVTFCEKEVDATAIDKE--IRGE 548  
 QY 639 GGTNSITVRNATFTWASDPPTLNGITFISIPGALVAVVQVCGCKSSLSALLAEMDKV 698  
 DB 549 LYTNTVEIHSGSFAWDSAEARILSDIEFLAGSKELVTVVSGVSGKSLLLAALGEMEKV 608  
 QY 699 EGHVAIKGSVAVVPOQAMTQNDLSRENILFGQLEPYPYRSVIOACALLPDLLEILPSGDR 758  
 DB 609 CGYVGRSVAVLSQPPHILNQLKXNVLMQADLNDVLYKKVIESCALKEDLKQLPDGDD 668  
 QY 759 TEIGKGVNLGGQKORVSLARAVYNSADYILFDDPLSAVDHVGKHFENYVGPGLMLK 818  
 DB 669 TEIGKGVNLGGQKARIALARAVYNSADYILFDDPLSAVDHVGKHFENYVGPGLMLK 728  
 QY 819 NKTRILVTHSMGYLPOVDVIVVMSGGKISEMGYQELIARDGAPAEFLTYASTEQEDA 878  
 DB 729 HTTRILVNTCTSFLOESGKIIVMKG-----ETRYESSG 761  
 QY 879 ENGVTGSGPGKEAKQWENGMLVTDGSAKQOLQRLSSSSSYSGDISRHNSHNSAELOKAE 938  
 DB 762 EESG-----GEENSIDLPGSIAGSRMSLSR--LSK-----ISRKSSSIVEK-- 804  
 QY 939 AKKEETWKLMEADKAQOQGVKSLVYDYMKAIGLFIISFLSIFLFCMCHVS--ALASNYWLS 997  
 DB 805 -KKPDA--LITKEBAAGRVNPGVLLYFKAMGIVTVYVLPVAIAVAVVNVVSVFALGRSLWT 861  
 QY 998 LWTDDPIVNGTQHTKV--RLSVYCALGISQGIAGVGYSMVSIIGILASRLCHLVDLHLS 1055  
 DB 862 AWSMDANIDINHPDPTMSVGARLVAGVAGFITEVIFLFLVLLLLIGVAAASKNLHPLLN 921

QY 1056 ILRSPMSFFERTSGNLVNRFSKELDTVDMSIPEVIMKFMGSLFNIVIGACIVILLATPTA 1115  
 Db 922 VLNPLSYFDITPIGRILNKLAKMEVVDURLSSFRFLVMAINM--TVLIVSYTTPLF 979  
 QY 1116 AIIIPPLGLIYFFQRYVYASSRQLKRLSVSRSPVYSHFNETLLGVSVIRAPAEQERFI 1175  
 Db 980 IAIIPFVFIYFFVLYKYSIKSTRQLOKIASLTRSPISFSETLQIGISTVRAPQWDEFV 1039  
 QY 1176 HQSLDKVDENOKAYPSIIVANRWLAVRLECVGNCIVLFAALFAVISRHS--LSAGLVGLSV 1234  
 Db 1040 RNDENHLNTHVKCSYYSQANRWLSIRLEGLNIVIFSAAILAIIGKESGITAGMLGLSV 1099  
 QY 1235 SYSLOVTVTLNMLVNRMSMETNIVAVRELKEYSETEKAPMOIQETAPSSWPQGRVE 1294  
 Db 1100 SYSLNITFMLNMFVQINEVETNVSVERIDEVSKTSAEWRLDNNIPSNWPTGGAVN 1159  
 QY 1295 FRNYCLRYREDDLVLRHNVITNGEKGIVGRTGAGKSSLTGLFRINESAEGEIIID 1354  
 Db 1160 IEDYSCRYRDELVLKQISLNLPGQKVGVCGRGTGAGKSSLALALFRIVEAADGNISID 1219  
 QY 1355 GINIAKIGLHDLRFKTIIPQDPVLFSGSLRWNLDPFSQYSDDEWVTSLEAHLKDFVSA 1414  
 Db 1220 QITSHIGHLDUREKTIIPQENVLPANTLRNIDPKQGTDOQLWALENSLKAHVBL 1279  
 QY 1415 LPDKLDHECAEGENLSVQORQLVCLARALLRKTKILVLDEATAAVDLTDDLIOTIRT 1474  
 Db 1280 LPHKLESVAEGENFSVQORQLLCTALLRKSVLVLDEATAGIDNFTDMVQATIRE 1339  
 QY 1475 QFEDCTVLTIAHRLTIMDYTRVVLVDKGEIQEYGAPSDLOOR--GLFFYSMAKDAGLV 1531  
 Db 1340 KPADSTIITIAHRLTIIDYDRIIVMDAGRIVEDGIPGELLKNRNSQFYGLAKSAKIV 1397  
 RESULT 10  
 ABC transporter SPAC3F10.11c [similarity] - fission yeast (schizosaccharomyces pombe)  
 N:Alternate names: metal resistance protein homolog  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C:Accession: T38712; T43244  
 R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
 submitted to the EMBL Data Library, February 1996  
 A:Reference number: Z21807  
 A:Accession: T38712  
 A:Status: translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-1478 <MUR>  
 A:Cross-references: UNIPROT:Q10185; EMBL:Z69369; PIDN:CAA93309.1; GSPDB:GN000066; SPDB:SE  
 A:Experimental source: strain 972H(-); cosmid c3F10  
 R:Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.  
 DNA Res. 4, 363-369, 1997  
 A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.  
 A:Reference number: Z17323; MUID:98162722; PMID:9501991  
 A:Accession: T43244  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: mRNA  
 A:Residues: 1043-1286, 'L', 1288-1300, 'Y', 1302, 'H', 1304-1310, 'H', 1312-1466, 'V', 1468-1478  
 A:Cross-references: EMBL:D89231; NID:g1749669; PIDN:BAAL3892.1; PID:g1749670  
 A:Experimental source: strain PR745  
 C:Genetics:  
 A:Gene: SPDB:SPAC3F10.11c  
 A:Map position: 1  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
 C:Keywords: ATP; transmembrane protein  
 Query Match 25.4%; Score 2548; DB 2; Length 1478;  
 Best Local Similarity 38.7%; Pred. No. 9.2e-163;  
 Matches 560; Conservative 285; Mismatches 509; Indels 92; Gaps 21;  
 QY 115 ITTILATFLIQLERRKGVQSSGIMLTFWLVA-LVICALAILRSKIMTALKEDAQVDFRDI 173  
 Db 96 IITYLALFLHVVEQPTLRIPMASLMEFLFKFLASALVLLLRPNYTF---PMLNVVPSI 152

QY 174 TFVYVPSLLLIQVLVS-----CFSDRSPLFSETHIDPNCPRESSASFLSRITFWITGLI 228  
 Db 153 TFFCSIVCLLAIYVPPANRWVYPPDAAELEST--GLRPSRTYANIFSRISFGWLSPLM 210  
 QY 229 VGYRQPLGSDMLINKEDTSEOVVPLVVKWKKECAKTRQPKVYVSSDKDPAQPKS 288  
 Db 211 KFGYRNYLTESDAWSLPPAERSNLTIVFEKNWISH-----AKKKKS 252  
 QY 289 SKVDANEVEEALIVKSPQKEMNPSLFKVLXKTFGVPFLMSPFKAIHDLMMFSGPOLKL 348  
 Db 253 SLV-----NWGVLFNLHKLTVVILVLQVDVAVFIQNLIRK 291  
 QY 349 LKLFVNDTKA--PDMOGYFVTVLLFVTAQLTDLVHLYFHCIFVSGMRKIKTAVIGAVY 404  
 Db 292 IVIFVSSYSSEHPQPPQVGFSLAIAMFLTNVVTALLOQYFQGLGMVGNWRSELTITAIY 351  
 QY 405 KRALVITNSARKSVTGEIVNLSVDAQRFMDLATYINWISAPLOVITLALYLLMLNLP 464  
 Db 352 KXSLLSAARQSRSGDINVMTMSVDTQKCDLTWFLFVIVSGPFQIVLALTNLVHLVGY 411  
 QY 465 SVLAGVAVMVLMPVNAVMAKTKTYQVAHMKSKDNRIKLMEILNGIKVLKLYAWELAF 524  
 Db 412 GALSAGVTFLLFPCNVVVIASIFKRFQNRQKNKQARSQFMTIINNRSIKLYAWENIF 471  
 QY 525 KDKVLAIHQ--BELKVLKKSAYLSAVGTTTWTCTPPLVALCTFAVVTI--DENNILDQA 582  
 Db 472 LOKLLQRLNTRLMKKGIVNTIGNFTLWFLFAPILVSAATFGTFIVLYGTRVLSVDIV 531  
 QY 583 FVSLALFNILFPNLPMTVISSIVQASVSLKRLRIPLSHBELEPDSITERPVKDDGGTN 642  
 Db 532 FACLSLFLNLOPFLTMFLPIVVSVLEASVALSRVYGLTAGELDSNAVQRIPANKEPSGV 591  
 QY 643 SITVRNATFTWA----RSDPPTLANGITFSIPEGALVAVVGVGCKSLLSALLAEMDKV 698  
 Db 592 CLEIKGTGFSWGPQNAAEPLRIDFVARRGELCCIVGKVGKMSLLEACLGNMQKH 651  
 QY 699 EGHVAIKGSVAVYVQQAQWIONSLRENTFCQLEEPYRVSIVQACALLPLEILPQSDR 758  
 Db 652 SGSVFRCSIIAYAAQQPWILNATQENILFLELDPEFYKTIIRACCLLRDFEILADGDQ 711  
 QY 759 TEIGKGVNLGGQKQVRVSLARAVYSNADYIFDDPLSADVDAHVGHKHFENVIKPKMLK 818  
 Db 712 TEVGEKGISLGGQKARISLARAVYSRSDIYLLDDILSAVDQHVNRDLVRNLLSGKLLR 771  
 QY 819 NKTRILVTHSNLYLPQVDVIVVMSGGKISEMGSVQELLAR--DGAPAEFLRYASTEQED 877  
 Db 772 SRCVILSTNSLTVLKEASMIYMLRNGKIIESGSPQLSSSPDSQLFOLLSEFSKDT--- 828  
 QY 878 AEENGVTGVSQPGKEAQOMENGLVTPDSAGKQLQRLSSSSSYSGDISRHH-----NSTA 932  
 Db 829 -----ASTGADTPLSRQSQVITSST-----DVTSSASRSDTVSNPKATIKGTG 874  
 QY 933 ELQKAEAKKSETWKL--MEADKAQTQGVKLSVYWDYMKAIGLFISPLSIFLFMCNHYSA- 989  
 Db 875 RIRK-RLTDEDNVKATQAAAEKMERGVKVKVYVYFKACSLFLIFL-YFLFIIGGIMN 932  
 QY 990 LASNYWLSLMTDDPIVNGTQEHKTVRLSVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1048  
 Db 933 VGINVWLKHWSEVNTQGVNPKPYFYLGITYLTLGLSALISLSSLTITVFCIAKSCRYL 992  
 QY 1049 HVDLHLSILRSPMSFFERTSGNLVNRFSKELDTVDMSIPEVIMKFMGSLFNIVIGACIVI 1108  
 Db 993 HDSNVKAVLRAPMSFFERTTGRILNFRFSDVYRVDEVISRFVWFFRNLFQIVFVLAVI 1052  
 QY 1109 LLATPTAIIIPPLGLIYFFQRYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1168  
 Db 1053 CYSPMPMILIVLFFLYRYNYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1112  
 QY 1169 EOEERFTHQSLDKVDENOKAYPSIIVANRWLAVRLECVGNCIVLFAALFAVIS--RHSL 1226  
 Db 1113 DMEDTFISENDIRVDNHRIFWFLYFSNRWQARVEAIGALVVFSSAFFGLSAVRGNPN 1172

QY 1227 AGLVGLSVSYSLQVTTYNLWLVMSSEMETNIVAVERLKEYSETEKEAPMOIQETAPSS 1286  
 DB 1173 SGLVGLSVSYAVQITQSLTFVVRQSDVETNIVSVERMLEYIGLPSEAPSIIPOHRPPEG 1232  
 QY 1287 WPOGVEFRNYCLRYREDLPVLHSHVNTINGGKVGIVGRTGAGKSSLTGLFRINES 1346  
 DB 1233 WPSHGAIFDHVSRYRENPLVNDISVNIKQEKIGIVGRTGAGKSTLTALFRLIEP 1292  
 QY 1347 ABGEIIDIIGINTAKTGLHDLRPKITIIPQDPVLPFSGSLRPMULDPPSQVSDREWTSLELA 1406  
 DB 1293 TSGDIQDDINITSIGLHDLRSRLAIIPQENQAFEGTIRENLDPNANATDEEIHWALEAA 1352  
 QY 1407 HLKDFVSALPDKLDEHCEGEGNSVGORQVLCARALLRTKILVLDEATAAVDLETD 1466  
 DB 1353 SLQFIQTLDDGLYRVTGEGANLSSGQQLMCLTRALLTPTRVLLDEATAAVDVEIDA 1412  
 QY 1467 LTQSTIRTFEDCTVLTIAHRLNTMDYTRVTVLDKGEIQEYAGPSDILQOR-GLFYDMA 1525  
 DB 1413 IVQRTIRERFNDRTILTIAHRLNTVMSNRILVLDHGKVFEDSTKLEENKASLFYSIA 1472  
 QY 1526 KDAGLV 1531  
 DB 1473 KESGLI 1478

RESULT 11  
 T01369  
 ABC transporter AtMRP2 [imported] - Arabidopsis thaliana  
 N:Alternate names: multidrug-resistance protein homolog T29F13.13  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
 C:Accession: T01369; D84759  
 R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
 submitted to the EMBL Data Library, May 1998  
 A:Description: Arabidopsis thaliana chromosome II BAC T29F13 genomic sequence.  
 A:Reference number: Z14179  
 A:Accession: T01369  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1623 <R0U>  
 A:Cross-references: UNIPROT:O64590; EMBL:AC003096; NID:g3132469; PID:g3132479  
 A:Experimental source: Cultivar Columbia  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
 eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: D84759  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1623 <STO>  
 A:Cross-references: GB:AE002093; NID:g3132479; PIDN:AAC16268.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g34660; T29F13.13  
 A:Map position: 2  
 A:Introns: 110/3; 170/2; 236/1; 274/2; 301/2; 309/3; 328/3; 350/3; 375/3; 411/3;  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
 F;1252-1452/Domain: ATP-binding cassette homology <ABC>

Query Match 23.8%; Score 2384.5; DB 2; Length 1623;  
 Best Local Similarity 35.2%; Pred. No. 1.1e-151;  
 Matches 561; Conservative 274; Mismatches 534; Indels 223; Gaps 30;  
 44 FYWLACPPF-YFVLSRHDGTYQMPPLNTKKTALG---FLMIVCWADLFYSFW----- 94  
 5 FIEWYKFPVNGVMTQVANAFGAYTFCATDSFVLGISQLVLLVLC-----LYRIWALXKD 60  
 95 -----ERSRGIFLA-----PVFLVPTLIGITTL-----LATFLIQLERRK 130  
 61 HKVERFCRLRLNYFLALLAAYATAEPLRL---INGISVLDFDGPGLPPF----- 109  
 131 GVOSGIMLTFWLVALVLCALAILRSKIMTALKEDAQVDLPRDITFYVYFS-----LL 182

DB 110 --EAFGLGVKAPANGAVVWMLMETKI-----YIRELRWYVFAVIALVGVDMV 156  
 QY 183 LIQVLVS-----CPSDRSPLFSETHO--- 204  
 DB 157 LLNLVLSVKRYSYVLYLYTSEVGAQVLFGLFEMHLPNLDTPYGVNPKVRSVETDIY 216  
 QY 205 -----ENPCPESSASFLSRTITFWITGLVIRGQPLEGSDLSLNKEDTSEQVVPVLV 258  
 DB 217 EETSDGOQIICPEKHANIFDKIFFSWMPLMTLSKRLTEKDVWYLDTDQTELTFTSFQ 276  
 QY 259 KWKKECAKTRKQPKVYVYSKQDPAQKSSKVDANEVEALIVKSQKWNPSLFKVLV 318  
 DB 277 HSWDKELQK-----PQ-----PWLRLALN 295  
 QY 319 KTFGPYFLMFFPKAIHDLMMFSGPOLKLLIKFVNDTKAPDMQGYFVTVLLFTYACLOT 378  
 DB 296 NSLGGRFWNGFWKIGNDQCFVGLLNLQKLSMQB-DAPAMNGYIYAFSIFGVVFGV 354  
 QY 379 LVLHQYFHCFSVGMRIKTAIVGAVYRKALVITNSARKSTVGBIVNLSMVDARFMDLA 438  
 DB 355 LCEAQYFQNVNRVGYRLRSALIAAFKSLRLTNEGRKFTQKITNLMTTDAESLQOIC 414  
 QY 439 TYINWISAPLOVILALYLLMLNLPSPVLGAVVWLVNPNVNAVMAMTKTYQVAHMKSK 498  
 DB 415 QSLTMSAPRIIILITLILYQQLGVASLIGALLLVLMFPLOTVIISKQKLTKEGLQRT 474  
 QY 499 DNRIKLMNEILNGIKVLKYAWELAFKDKVLAIRQELKVLKKSAYLSAVGTFVTVCPFF 558  
 DB 475 DKRIGLMNEVLAAMDVTAKCYAWENSFGKQVTVDDLSWFRKSKQLLGNLNFILSPV 534  
 QY 559 LVALCTFAVYVITDINNILDQAFAVSLALFNILRPLNLPIMVSISSVQASVSKRL-- 616  
 DB 535 LVTIVSGVFTLLGDD--LTPARAFTSLSLFAVLRFLPLFPLNIIITQVNVANVSLKLEE 592  
 QY 617 -----RIFLSHEELEPPDSIERPVDGCGGTSITVRNATFTW-ARSDPTLNGITFSIP 669  
 DB 593 VLATEERILNPPPIEP-----GPAISIRNGYFSWDSKGRDPTLSNLINLNDVP 640  
 QY 670 EGALVAVVGVQVCGKSSLSALLAEMDKV-BGHVAIKGSVAVVPOQAIONDSLRENILF 728  
 DB 641 LGLSLVAVVGVSTGEGKTSLSLAILGELPATSDAIVTLRGSVAVVPOQVSIENATVRDNL 700  
 QY 729 GCQLEPYRVSIVQACALLPDLLEILPSGDRTEIGKGNLGGQKQVSLARAVYSNADI 788  
 DB 701 GSPDREKYEIRADIVTSLKHLLELPLGGLDLTEIGERNVIGSGGQKQVSMARAVYSN 760  
 QY 789 YLFDPLSAVDHGVCKHIFENVIGPKMLKNKTRILVTHSMSTYLPQVDVIIVMSGKISE 848  
 DB 761 YIFDDPLSALDAHVGQGVFEKCI--KRELQKTRVLVTNQLHFLSQVDRIVLVHSGTVKE 818  
 QY 849 MGSYQELLARDCAFAEFLRTYASTEQEQDAEENGVTGSGPKE--AKOMENGLMLVDTSA 906  
 DB 819 EGTVEELSSNGPLFOLMENAGKVEEY--SEENGAEADQTAEQVPAVNGNTNGLQMGSD 876  
 QY 907 GKQORQSSSSSSSGDISRHHNSTAELOKAKAEETWKLMEADKAQTOGVKLSVYWDY 966  
 DB 877 DKK-----SKEGNKGGKSVLIKQBERETGVVSVWRVLRKY 911  
 QY 967 MKAI-GLFISFLSIFLFCNVHVSALASNVLSLWTDTPVNGT-QEHTRVRLS-VYGALG 1023  
 DB 912 QDALGAGVWVMMLLCYLVTVEVFRVTSSTLSEWTD----AGTPKSHGPLFYNLYALLS 967  
 QY 1024 ISQGIAGVFGYSMAVIGGILASRCLHDLHLSILRSPMSFFBRTPSGNLVNRFSELOTV 1083  
 DB 968 FQGVLTLTNSYMLMSSLSYAAKLDHNLHLSILRAPSFFHTNPLGRINRPAKDLGDI 1027  
 QY 1084 DMIPEVIMKMGSLFNIVGACIVILLATPIAAIIPPLGLIYFFQVRYVASSRLKRL 1143  
 DB 1028 DRTAVFVNMFMFGVQSLLSTVYLVGIVSTLSLWAIMPLLVLFYAGLYIQTAREVRKM 1087  
 QY 1144 ESVSRSPVYSHNETLLGVSVIRAFEEQERFTHQSDLKVDENQKAYPYSIVANRWLAVRL 1203

1088 DSISRPVFAQGEALNGLSTIRAYKAYDRMADNGRSMNNIRFTLVNMGANRWLGIRL 1147  
 1204 ECVCNCIVLPAALFAVI-----SRHSLSAGLVGLSVSYSLQVTTYLNWLVMSSEMETNI 1258  
 1148 ETGLGLMIWLTASFVWONGRAENQOAFATMGLLLSYALNITSLLTGVRLRLASLAENSL 1207  
 1259 VAVERLKEYSETEKEAPWOIQTETAPPSSPOVGRVFNVCYRREDLPVLRHINVTIN 1318  
 1208 NAVERVGNVIEIPEAPPVIEENRPPGPPSSGSIKFEDVULRYRQPLPPVILHGVSFFIH 1267  
 1319 GGEKVGIVGRTGAGKSSLTGLPRINESAGEHIIIDGINIAKIGLHDFKTIITIPQDPV 1378  
 1268 PTDKVGIVGRTGAGKSSILNALPRIVEVEKGRILIDDGCVGKFGMLDKRLKVLGIIPQSPV 1327  
 1379 LFSGSLRMNLDPESQYSDREVWTSLELAHLKDFVSNALPKLDHECAGGENISVQGRQIVL 1438  
 1328 LFSGVTRFLNDPPEGHNDADLESERAHLKOTIRRNPLGLDAEVSAGEENFSVQGRQLL 1387  
 1439 CLARALLRKTILVLDDEATAAVIDTDDLIQSTIRTOFEDCTVLITTAHRLNTIMDYTRVI 1498  
 1388 SLRALLRRSKILVLDDEATAAVIDVTRDALIQTIREEFKSCITMLIAHRLNTIIDCDKIL 1447  
 1499 VLDKGIQYGAUSDLLQGRGLFYS-MAKDAG 1529  
 1448 VLDSCRVOBFSSPENLLSNEGSSFSKMWQSTG 1479  
 RESULT 12  
 D86428  
 Glutathione S-conjugate transporting ATPase (AtMRP1) - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: D86428  
 R:Theologian, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: D86428  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1622 <STO>  
 A:Cross-references: UNIPROT:Q9C8G9; GB:AE005172; NID:g11055814; PIDN:AAG28284.1; GSPDB:C  
 C:Genetics:  
 A:Map position: 1  
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
 Query Match 23.5%; Score 2356.5; DB 2; Length 1622;  
 Best Local Similarity 35.6%; Pred. No. 8.6e-150;  
 Matches 569; Conservative 274; Mismatches 564; Indels 193; Gaps 31;  
 47 WACFPF-YFLYLSRHDGVIQMTPLNKTALG-----FLLMIVCWADLFYSEW 94  
 8 WYCKVPNGVTKTVDAFGATPCADSFVLGISHLVLLILCYRLWLIITKHDKVDFC 67  
 95 ERSR--GIFLA-----PVFLVSPTLTGITTL-----LATFLIQLERRKGVQS 134  
 68 LRSKWFYFLAALAYATAEPLFRL--VMSRISVLDDGAGFPPEAFMLVLE--AFAWG 122  
 135 SGIMLT-----WLVALCALAILRSKIWTALKEDAQVDLFRDITFYVYFSLLL 183  
 123 SALVWTVETKTYIHELKRYRFAVIVALVGDVMLNLVSVK-EYVGSFKLYIYISEVA 181  
 184 IQLV-----LSCFSDRSPLFSETIHD-----PNPCPSSASFLSRIFFW 223  
 182 VQVAFGTLFVVFNNLDPPGYTPVGTENSEDYEEELPGENICPERHANLFDSEIFSW 241

224 ITGLIVRGYROPLEGSDLWSLNKEDTSEQVVPVLVKNWKKCECAKTRKQPVKVYSSKOPA 283  
 242 LNPLMTLGSKRPLTEKWHLDTDWKTETLRSFKQSWDKELEK----- 285  
 284 QPKSESSVDANEVEEALIVKSPQKWNPSLFKVLKTPGYPFLMSFFFKAIHDLMMFSGP 343  
 286 -PK-----PWLRLANNLSLGGFRWGMGFKIGNDCSQFVGP 320  
 344 QILKLLIKFVNDTKAPDWQGYFYTVLLFVTACLOTVLHVOYFHCFSVGMRIKTAIVGAV 403  
 321 LLLNELLKSM-QLNEPAMIGIYIYALSFVGVVLGVLCEAOYFQNVMRVGYELRSALIAV 379  
 404 YRKALVITNSARKSTVEGEIVNLSVDAQRPMIDATYINMTWSAPLOVILALYLWNLG 463  
 380 FRKSLRLTNEGRKKFQTKITNLTMTDAESLQQICQSLHTMTWSAPFRIIIVLVALVLYQOLG 439  
 464 PSVLGAVAVMVLVPMVNAVMAMKTKYQVAMKSKDNRIKLMNEILANGIKVLKLYAWELA 523  
 440 VASIIIGALFLVLMFPIQTVIISKTKLTKEGLOTKRIGLMNEVLAAMDVTVCYAWENS 499  
 524 PKDKVLAIROBELKVLKKSAYLSAVGTFTWCTPFLVALCTFAVYVITIDENNILDATAF 583  
 500 PQSKVQTVRDDDELQSWFRKAQLLSAFNMFILNSIPVLTVVSVFVSLGGD--LTPARAF 557  
 584 VSLALENLIREPLNIPMWISSIVQASVSLKRLRIFLSHEE--LEPDSIERRPVKDGGT 641  
 558 TSLSLFSLVLRPFLPMLPNIITQWNVANVSLNRLVEEVLSTEERVLNPN---PIEP--GQ 611  
 642 NSITVRNATFW-ARSDPPTLNGITFSIPEGALVAVVGVGVCGKSSLSALLAEMD-KVE 699  
 612 PAISIRNGYFSWDSKADRPILSNLNDIPGLSLVAVVGVSTGEGKTSLSIAMLGELPARSD 671  
 700 GHVAKGSVAYVPOQAWIQNDSURENLFGCOLLEPYRVSIVQACALLPDLLELPDSDRT 759  
 672 ATVTLRGSVAYVPOQSWIFNATVRDNLIFGAPFDQEKYERVIDVTALQHDLELLPGDLT 731  
 760 EIGSGVNLSSGQKORVSLARAVVSNADIVLFDPLSADVAHVGHKHFENVIGPKMLKN 819  
 732 EIGRGNVSGGQKQVSMARAVVSNDSVCLDDPLSALDAHVGGQVFEKCI--KRELQ 789  
 820 KTRILVTHMSYLPQVDVIVISGGKISEMGYSQELLDARGAFAEFLRTYASTEQEQDAE 879  
 790 TTRVLVTNQLHFLSQVDKILLVHEGTVKEBGTVEELCHSGPLFQRLMENAGKVEDY--SE 847  
 880 ENGVTGVSFGPKKAKQWENGLVTDNAGKQLQRLSSSSSYSGDISRHNSHSTAELOK--- 936  
 848 ENGAEVD--QTSVKPVENG-----NANNLQKXGI 875  
 937 -AEAKKEETWKLMEADKAQGVKLSVYWDYMKAI-GLFTISFLSIFLFCMCHVYSALASNY 994  
 876 ETXNSKEGNSVLKREERETGVSVKVLERYQNALGAWVMMLVICYVLTQVFRSSST 935  
 995 WLSLWTD--PIVNGTOEHTKVRLSVYGALGISQGIADVFGYSMAVSIIGGILASRCLHVDL 1052  
 936 WLSWTDGSGPKTHGFLFYN---IVVALLSFGQSVTLINSYVLMSSLSLYAAKQMDAM 991  
 1053 LHSILSPKSPFEERTPSGNLVNRFKSELDTVDSMIPEVIMKFMGSLFNVIACIVILLAT 1112  
 992 LGSILKAPWVFFOTNPLGRILINEFAKMDGIDITVAVFVNMFMGSAQLLSTVILIGIVS 1051  
 1113 PIAAIIIPPLGLIYFFQVRFYVASSRQLKRLSESVSPVYSHENETLLGVSVIRAFEEQE 1172  
 1052 TLSLWALMPLLVVYFAYLYQNTSREIKEMDSTTSPVYQAFGEALNGLSIRAYKAYD 1111  
 1173 RPTHQSDLKVDENQKAYYSIVANRWLVRELCVGNICVLFALFAVI-----SRHLSA 1227  
 1112 RMAENGRSDMNNIRFTLVNMAANRWLGIRLEVLGGLMVLTLASLWQMONGKAANOQAVA 1171  
 1228 GLVGLSVSYSLQVTTYLNWLVMSSEMETNIIVAVERLKEYSETEKEAPQIOETAPSSW 1287  
 1172 STWGLLSVALSTSSLTAVRLASLAENSLNSVERGVNIIPIPSAPLVIENRRPPGW 1231  
 1288 PQVGRVEFRNYCLVRREDLDFVLRHINVTINGEKEVIGVTRTGAGKSSUTLGLFRINESA 1347

```

1232 PSKSGIKFEDVVLRYRPELPPVLLHGVSLISPMKVGIVGRTGACKSSLLNALFRIVELE 1291
1348 EGEIIDDGINIAKIGLHDLRFKTIIPQDPVPLFSGSLRWNLDPFSSOEVEWTSLELAH 1407
1292 KERILIDECIDGRFLMDLRKVLGIIPQAPVPLFSGVFRNLDPFSSHNDADWESLERAH 1351
1408 LKDFVSALPKLDHCEAGENLSVGORQLVCLARALLRKTILVLDEATAAVIDETDOL 1467
1352 LKDTIRNPLGLDAEVEAGENFSVGORQLLSARALLRSKILVLDEATAAVIDRTDVL 1411
1468 IOSTIRTOEDCTVLTIAHRLNTIMDYTRIVLVDKGEIOEYCAPSDLLQOGLFYFSMAKD 1527
1412 IOKTIREEPKSCMTLIIAHLNTIIDCDKVLVDGKGVQSPSPENLLSNGESSFSK--- 1468
1528 AGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAEPYRTRE 1567
1469 -----MWQSTGANAAYLRSITLENK-----RTR 1493

RESULT 13
E86428
probable ABC transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: E86428
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
aneen, N.P.; Hughes, B.; Huiziar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86428
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1495 <STO>
A:Cross-references: UNIPROT:Q9C8H0; GB:AE005172; NID:g11055818; PIDN:AG28288.1; GSPDB:G
C:Genetics:
C:Map position: 1
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 23.2%; Score 2326; DB 2; Length 1495;
Best Local Similarity 34.3%; Pred. No. 8.6e-148;
Matches 541; Conservative 314; Mismatches 566; Indels 154; Gaps 28;

QY 2 ALRGFCS--ADGSDPLMDNNTWNTSNPDTKCFQNTVLVWVP-----CFY-LWACPP 51
DB 5 ALNWKCPVADGF-----WEKAVDGAFGAYTCADSLVMLVSHFVLLGLCFYRWIIF- 58
QY 52 FYFLYLSRHDRGYQWTPNPKTKTALGFLMLWICWAD-----LPYSFWRSGRGIPLAPV 105
DB 59 -----HNTKAIQYVLRKKNYCNVLG-LLACVYVPEVLRVLMGSLPDMDEETDFPFP 110
QY 106 FLVSPILLGITLTALLFLQLERRKQVQSSGIMLFWLVALVCAALILRSKMTALKEDA 165
DB 111 EVASLMVEFAFWMSMLVLGLTETKQYKFRVYRFGVLVLAVALDVLPLKNSIN 170
QY 166 QVDLFRDIT---FYVYFSLLLQLVLSCFSDRSPIFSETHDP-----NPGP 209
DB 171 RTALYLFISRCSQALFGILLIYIPEL--DYPFGHIVNNEPLDNEVDALRGEGHICP 228
QY 210 ESSASFLSRTFWITGLIVRGQPLEGDLWSLNKEDTSQGVVPLVKNWKBCKATR 269
DB 229 ERHASIFSRIFYGWITPLMQLGKRPITEKDVQWLDKWDQETTLIKRFQRCWTESSRR-- 286
QY 270 KQPVKVVYSSKQPAQKSSKVDANEEVALIVKSPQKEWNSLFLVLYKTFGPFMLMSF 329
DB 287 -----PK-----FWLLRALNNSLGRFWLAG 307

```

```

330 FFKAIHDLMMFSSPOILKLLIKFVNDTKAPDMQGYFYTVLLFVTAQLOTLVLHQYPHICF 389
308 IFKIGNDLSQFVGPVILSHLLRSQMGGD-PAWGVYVAFIIFVGVTLGLVCEAQIFQNVW 366
390 VSGMIKTAIVAGAVRKALVITNSARKSTSTVGEIVNLMSVDAQRFMDLATYTNMWSAPL 449
367 RVGPFRLRSTLVAIIFFKSLRLTHEARKNFASPKVTNMITTDANALQQIISQOLHGLMSAPP 426
450 QVILALYLMLNLGFSVLGAVVAVMLVPVNAVMAKTKTYQVAHMKSKDNRIKLMNEIL 509
427 RIIVSMILLYOQLGVASLFGSLILFELIPLQTLIISKRKLTKEGLOWTDKRVGITEIL 486
510 NGIKVLYAWELAFKDKYLAIRQBELKVLKKSAYLSAVGTFTWCTPFLVALCTFAYVY 569
487 SSMDTVKCYAWEKSPESRIGIRNEELSWFRKAQLLSAFNSFILNSIPVVTWVSGVVF 546
570 TIDENNIIDQAOTAFVSLALFNILRPILPMLPWISSIIVQASVSLKRL-RIFLSHELEPDP 528
547 LLGGD--LTPAPRAFTSLSLFAVLRPLNMLPNLLSQVNVNANYSLQRIEELLSEERILAQ 604
629 STERRPVKDGCGTNSITVRNATFTW-ARSDPPTLNGITFSPIDEGALVAVGVGVGCKSSL 687
605 NPPLQ-----GTPAISIXNGYFSDSKTKTPTLSDINLEIPVGLVAVGVTGEGKTSL 659
688 LSLALAEKDKVE-GHVAIKGSVAYVPOQAQIOWNSLRENILFGCQLEEPYRSVIOACAL 746
660 ISAMGELSHAETTSVWIRGSAVVPQVSWIFNATVRNENILFGSDFESERYWRAIDATAL 719
747 LPDLBILSPGDRTEIGEGKVNLSGQKQKORVSLARAVYSNADYLFDDPLSAVDHVGKHI 806
720 QHDLPLFGRLDTEIGERGVNISGQKQKORVSNARAVYSNDSYVIFDDPLSALDAHVAHV 779
807 FENVIGPGKMLKNKTRILVTHSMSVLPQVDVIVMVGSGKISMGVSQOELLARDGAFABFL 866
780 FDSCH--KDELRGKTRVLVTNQLHFLPLMDKILVSEGMIKEEFTVELSKSGILFKKUM 837
867 RTYASTEQEQDAEENGVTGVSQPGKEAKQEMENGLVTPDSAGKQLQRLSSSSSYSGDISR 926
838 ENAGKMDATQEVNTN-----DENILKLGPTVTVDVS---ERNLGSK----- 876
927 HINSTAELOKABKEETWKLMEADKAQTQGVKLSVYNDYMKAI-GLFTSIFSLIFLPMCN 985
877 -----QKERRS-VLIQKBERETGIISWNLARYKEAVGGLWVVMILLACYLAT 924
986 HVSALASNVLSLWTDPIVNGTQEHK-----VRLSVYVYALGIGSQIAGVGYSMASVIG 1040
925 EVLRVSSSTWLSITWD-----QTSKNYSPGFYIVVYVALLGFGQVAVFTNSFWLITS 977
1041 GILASRCLHVDILLHSILRSPMSFFERTPSGNLVNRFPSKELDTVDSDMPEVIMFVMSLFN 1100
978 SILHAARLHDAMLSSILRAPMLFFHTNPTGRVINRFSKIDIGIDIRNVANLMMFMNLQMQ 1037
1101 VIGACIVILLATPIAAIIPPLGLIYFVQRPYVASSRQLKRLSVSRSPVSHFNETLL 1160
1038 LLSTPALICTVSTISLWAIMPLLIIFAYAYLYQSTSRVRLDSVTSRPIYAQGEALN 1097
1161 GVSIVRAFEEOERFIHQSDLKVDENOKAYPSIVANRMLAVRLEVCNGCIVLPAALFAYI 1220
1098 GLSSIRAYKADYRMAKINGKSMNNIRFTLANTSSNRMLTIRLETIGGMVIMLTATFVIL 1157
1221 SRHSLS-----AGLVGLSVYSLOVTTYLNMLVRMSSEMETNIVAVERKEYSETEKAP 1275
1158 QNGTNNQAGFASWTGLLSLTYLNTLSLGLVLRQASRAENSLNSVERVGNVLDLPSEAT 1217
1276 WOIQETAPPSNWPQVRVEFRNYCLURYREDLDVLRHINVTINGEKVIGVORTGAGKSS 1335
1218 DIENNRVPCWPGWPGSGSKFEDVHLRPLGPPVLLHGLTFFVSPSEKVGWVGRTGAGKSS 1277
1336 LTLGLFRINESAGEIIDIENIAKIGLHDLRFKTIIPQDPVPLFSGSLRWNLDPFSSQVYS 1395
1278 MLNALFRIVEKEGRIMIDDCDVAKFGTUDVRRVLSIIPQSPVLFSGTVRFNIDPFSEHN 1337

```

Qy 1396 DEEVWTSLELAHLKDFVSALPDKLDHCEAEGEENLVSQGOROLVCIARALLRKTILVLDE 1455  
 Db 1398 DAGLWEALHRAHIKDVISRNPFGLDAEVCEGENFSVQGORQLLSARALLRRSKILVLDE 1397  
 Qy 1456 ATAAVDLETDLDIOSTIRTPQEDCTVLITIAHRLTIMDYTRVILVDKGEIOEYCAPSDLL 1515  
 Db 1398 ATASVDVRTSLIORTIREEPKSTMLVIAHRLTIIDCDKILVSSGQVLEYDPORELL 1457  
 Qy 1516 -QQRGLFYSMAKQDG 1529  
 Db 1458 SRDTSAFFRMVHSTG 1472

RESULT 14  
 T27408  
 hypothetical protein Y75B8A.26 - *Caenorhabditis elegans*  
 C/Species: *Caenorhabditis elegans*  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C/Accession: T27408  
 R:Barlow, K.  
 submitted to the EMBL Data Library, November 1998  
 A/Reference number: Z20361  
 A/Accession: T27408  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-1144 <WIL>  
 A/Cross-references: UNIPROT:Q9XW61; EMBL:AL033514; NID:e1343251; PIDN:CAA22110.1; CESP:Y  
 A/Experimental source: clone Y75B8A  
 C/Genetics:  
 A/Gene: CESP:Y75B8A.26  
 A/Introns: 63/3; 120/3; 222/3; 271/3; 372/3; 451/1; 636/1; 687/3; 778/3; 953/2; 1042/2;

Query Match	23.1%;	Score 2312;	DB 2;	Length 1144;
Best Local Similarity	42.5%;	Pred. No. 4.8e-147;		
Matches	491;	Conservative 228;	Mismatches 381;	Indels 54; Gaps 16;
Qy	396	KTAVIGAVYKALVITNSAKSSKTVGEIVNLSVDAQRFNDLATYINMTWSAPLOVILAL	455	
Db	9	QTVLSNALHKKURLSPSARSNRTAGELNHAADVDEIIVHSVPYQLQNMWMSVPFQVTLAM	68	
Qy	456	YLLNLNLGPSVLAGVAVVWLVMPVNAVMAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVL	515	
Db	69	TMLAITLGMAMAGVCIMILFIPNLNCTSRFIKLSQKQMKIKDERTKLSNEMLNGIKVV	128	
Qy	516	KLYAWELAPKDKVLATROEBELKVKLSAYLSAVGTTFTWCTPPLVALCTPAVV--TIDE	573	
Db	129	KLYAWBESFEDQNRURAKREKVMRLNVCILSRIVDVANAASPELVAIGSFCTCVLWSPDE	188	
Qy	574	NNILDAQTVSLALFNILRFPNLPMWISSVQASVSLKRLIRIFLSHEELPDSIER	633	
Db	189	NG-LTSPVAFVALTIFNQLRQPRMNVANLINTLVQARVSNKRLRQFLNDEEM- ---RKT	243	
Qy	634	PVKDGGGTNSITVRNATFTW--ARSDPPTLNGITFSIPEGALVAVVGQVCGKSSLLSALL	692	
Db	244	EVALG--NAIVFKNASLWNKGPONPEVLKDLSATIKPGQLIAIVGSGGKSSLLSAVL	300	
Qy	693	AENDKVEGHVAIKGSVAVYPOQAWIQNDSLENILFGCQLEBYPYRSVTOACALLPDEI	752	
Db	301	DEWLIDGRVKVGGSIAYVPEHSMIFNKTIKENILFGNELSNFYFDYVWGSCOLKTDFRH	360	
Qy	753	LPSGDRTEIGEKGWNLGGOKQVSLARAVYSNADIVLPDDPLSAVDHVGKHI FENVIG	812	
Db	361	FQOQENTWVGENTILSGGOKARISLARAVYQDKDIYLLDDPLSAVDHVGFRALPKVIG	420	
Qy	813	PKGMLKNKTRILVTHSMYILPOVDVIVMSGGKISEMGYSQELLARDGAPAEFLRTYAST	872	
Db	421	PDGLLSKSRVLVTHNLQYTKYVDTTVIYEDGQIVQHGSPEDIAYVDGPFGR-LWSECE	479	
Qy	873	EQQDAEENGVTGVSFGPKAEKQWENGLVTSQAGKQLQRLSSSSSYSGDLSRHHNSTA	932	
Db	480	SDEVDAAEASSEASVTTPVPVPLENG---DNGAIEKSDIRTNHSPSEKSR-----	529	
Qy	933	ELOKAEAKKEETWKLMEADKAQTQGVKLVSYYMDYMKAILGLFISFLSFIIFMCHNVSALAS	992	

Db 530 ---KSEEPQKVEKNVE--NVQLGRVKSVYQLYIKTMGIENSSAFLFFIAHFTVMIIR 584  
 QY 993 NYWLSLWTDG-----PIVNGTQE-----HFKVRLSVYCALGISOGIAVFGYSM 1035  
 Db 595 SLWLSDDSNENAAIKKATLSVSDYLNSTSSVDGPVSVETRIUVVAGFG--GLEMLLLAL 641  
 QY 1036 A---VSGIILASRLCHVDLLHSILRSPMGSPFFETPSGNLVNRFSKELDTVDSMIPEVIK 1092  
 Db 642 AFTVLITIGSLRASVGLHSLPIHALLAVAPISFFDTPGRIINRLSRDLDVDK-LQDNIR 700  
 QY 1093 MFMSGLFNVICACIVILLATPIAAIIIPPLGLIYFFVQRFVVAASSRQLKRLSVSRSPVY 1152  
 Db 701 MCTQTLNACMLVLVLSISTPIFVCAAPLILYIFVMYIYPTSRQLKRLSANSRSPIL 760  
 QY 1153 SHFNETLLGVSVIRAFPEQERFIHOSDLKVDENOKAYYPSIVANRWLAVRLCEVGCIVL 1212  
 Db 761 STIAESIHGASSIRAFPKTERTTTALTSTNDVKFAQCYLSHMSNRWLATRLLELGNTCVL 820  
 QY 1213 PAALFAVIR--HSLSAGLVGLSVSYSLQVTTYLNWLVRMSSEMETNIVAVERLKEYSET 1270  
 Db 821 PASLSATLSKTYFGLTQPMAGLSVSYALTITEVLNLCVRGSVEIESNIVSVRNEYOKL 880  
 QY 1271 EKEAPMOIQTAP--PSSMPQGVRFEPNRYCLRYREDLDVLRHNVNTINGEKVGIVRT 1329  
 Db 881 EPEAPWRIEKSLNENKWPVKGIKELDFGSMRYKNLPVLVKNIDLLKIEGGERIGVIGRT 940  
 QY 1330 GAGKSSITLGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRNLMD 1389  
 Db 941 GSGKSSLTALYRMIEGESGTIKIDVDEITDGLHQLRSKLIIPQEVFVSGTLRFNLMD 1000  
 QY 1390 PFSOYSDEEVWTSLELAHLKDFVSALPKLDHCEAEGENLSQOROLVCLAFALLRKT 1449  
 Db 1001 PFNOYSDQIWNCLIEICQLQFOAEDKTLDRVIAEGGKNMSVGERQLCLCALLRGAR 1060  
 QY 1450 ILVLDEATAVLDLTDLLIOSTIRTPQEDCTVLTIARHLNTIMDYTRVIVLDKGEIOEYG 1509  
 Db 1061 IVILDEATASVDTVDGIVQRAIRQHPQSTTISIAHRLDTIVDSRIVVLDAGRVAEFD 1120  
 QY 1510 APSDLLQOGLFYS 1523  
 Db 1121 TPSNLLNPPDSLYS 1134  
 RESULT 15  
 S64757  
 Probable membrane protein YL015w - yeast (Saccharomyces cerevisiae)  
 N;Alternate names: hypothetical protein A255; hypothetical protein L1313  
 C;Species: Saccharomyces cerevisiae  
 C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 09-Jul-2004  
 A;Accession: S64757; S64763; PS0041; S69391; S70560  
 R;Micosg, T.; Zimmermann, F.K.  
 submitted to the Protein Sequence Database, May 1996  
 A;Reference number: S64743  
 A;Accession: S64757  
 A;Molecule type: DNA  
 A;Residues: 1-1559 <MIO>  
 A;Cross-references: UNIPROT:P14772; EMBL:Z73120; NID:g1360184; PIDN:CAA97460.1; GDB:1559  
 A;Experimental source: strain S288C  
 R;Goffeau, A.; Purnelle, B.  
 submitted to the Protein Sequence Database, May 1996  
 A;Reference number: S64761  
 A;Accession: S64763  
 A;Molecule type: DNA  
 A;Residues: 1-1559 <GOF>  
 A;Cross-references: EMBL:Z73120; NID:g1360184; PIDN:CAA97460.1; PID:g1360185; GDB:1559  
 A;Experimental source: strain S288C  
 R;Boy-Marcotte, B.; Damak, F.; Camonis, J.; Garreau, H.; Jacquet, M.  
 Gene 77, 21-30, 1989  
 A;Title: The C-terminal part of a gene partially homologous to CDC25 gene suppressor  
 A;Reference number: PS0041; MUID:89306677; PMID:2545538  
 A;Accession: PS0041  
 A;Molecule type: DNA

A:Residues: 1-255 <BOY>  
A>Note: the authors translated the codon CAG for residue 248 as His  
R:Purnelle, B.; Goffeau, A.  
submitted to the EMBL Data Library, April 1996  
A:Description: The sequence of 32 kb on the left arm of yeast chromosome XII reveals 14  
family and a new ABC transporter homologous to the human multidrug resistance protein.  
A:Reference number: S69380  
A:Accession: S69391  
A:Molecule type: DNA  
A:Residues: 1-1559 <PUR>  
A:Cross-references: EMBL:X97560; NID:g1297003; PIDN:CAA66162.1; PID:g1297015  
R:Miesga, T.; Zimmermann, F.K.  
Yeast 12, 693-708, 1996  
A:Title: Sequence analysis of the CEN12 region of Saccharomyces cerevisiae on a 43.7 kb  
e conductance regulator protein CFTR.  
A:Reference number: S70557; MUID:96405918; PMID:8810043  
A:Accession: S70560  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1559 <MIW>  
A:Cross-references: EMBL:X91488; NID:g1495203; PIDN:CAA62776.1; PID:g1495208  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995  
C:Genetics:  
A:Gene: SGD:BPT1; MIPS:YLI015W  
A:Cross-references: SGD:S0003938  
A:Map position: 12L  
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology  
C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein  
F:28-44/Domain: transmembrane #status predicted <TM1>  
F:143-159/Domain: transmembrane #status predicted <TM2>  
F:178-194/Domain: transmembrane #status predicted <TM3>  
F:334-350/Domain: transmembrane #status predicted <TM4>  
F:421-437/Domain: transmembrane #status predicted <TM5>  
F:526-542/Domain: transmembrane #status predicted <TM6>  
F:550-566/Domain: transmembrane #status predicted <TM7>  
F:654-847/Region: ATP-binding cassette homology <ABC1>  
F:672-679/Region: nucleotide-binding motif A (P-loop)  
F:974-990/Domain: transmembrane #status predicted <TM8>  
F:1017-1033/Domain: transmembrane #status predicted <TM9>  
F:1099-1115/Domain: transmembrane #status predicted <TM10>  
F:1118-1134/Domain: transmembrane #status predicted <TM11>  
F:1212-1228/Domain: transmembrane #status predicted <TM12>  
F:1319-1529/Domain: ATP-binding cassette homology <ABC2>  
F:1336-1343/Region: nucleotide-binding motif A (P-loop)

Query Match 22.9%; Score 2294.5; DB 1; Length 1559;  
Best Local Similarity 35.1%; Pred. No. 1.2e-145;  
Matches 574; Conservative 292; Mismatches 541; Indels 227; Gaps 39;

QY	23	NTSNPDFTKCFQNTVLVWVPCRYLW-ACPPFYFLYLSRHRDGYIOMTPLNKTALGFL-	80
DB	23	NALNP-----CFISVISAWQAVFLLIGSVQLWKLKYNKVPFPKFPPLPSKINSRHLT	78
QY	81	-LWIVCW-----ADLFYSFWMERSGIFLAPVFLVSPTLGI---TTLAT	121
DB	79	HLTNVCFQSTLIICELALVQSQSDRVYFPLK-KALYLNLF-----NLGISLPTQYLA	132
QY	122	FLIQLEKRGVQSGLMTFWLVALVCALAILRSKI-----MTALKEDAQV	167
DB	133	F-----KSTFSGMGNLFYMFQILLQLFLILQRYHGSSNRLTVISGQTAMILEVL	185
QY	168	DLPRDITFYVYFSLILLQLVLSFCFSRPLSFSTIHDNPNCPSSASFSLRITFWMITGL	227
DB	186	-LNSVAIFY-DLCIFEPI-----NELSEYKKNQWVP-----PVHVLSYITFIWNLK	233
QY	228	IVRGYRQPLEGSDIWSLUNK---EDTSEQVVPVLVKNWKKCAKTRKOPKVVVSSKDAQP	285
DB	234	IVETYR-----NKKIKDPNQLPPEVDLNK-----	259
QY	286	KESKVDANEVEALIVKSPQKWNPSLKVLYKTFGPFLMSPFFKAIHDLMMFSGPQI	345
DB	260	SISKEFKANWELEKWLNRN-----SLWRAIWKSFGRITISVAMLYETTSDLLSVVQPF	312

QY	346	LKLLIKFVN---DTKAPDMQGYFYTVLLFVTAQTQLVLVHQYHFHCFVSGMRKTAIVGA	402
DB	313	LRIFDGLNPETSKYPPPLNGVFIALTLFVISVVSFLTNQFYGIPEAGLIRGSLASL	372
QY	403	VYRKALVITNSAKSSTVGIVNLMVDA---ORFMDLATYINMWSAPLOVIALALYLW	459
DB	373	VYQKSLRLTLAERNEKSTGDLNLSVDVLIQIRFFENA---QTIIGAPIQIIVLVLSLY	429
QY	460	LNLGPSVLGAVVAVVLMVVPVNAVMAKTKYQVAHMKSKDNRIKLMNEILNGIKVLKLYA	519
DB	430	WLGLGAVIGLVLVTAIMPINAFLSRKVKLTKQMYKDKMRKTIITELLNAIKSLYA	489
QY	520	WELAFKDKVLAIRQE-ELKVLKKSAYLSAVGTFTWCTPFLVALCTFAVVTIDENNILD	578
DB	490	WEEPMARLNRHVRNDMELNFKRIGIVSNLIYFANNCVPLMVTCTGTLF-SLFSDSL	548
QY	579	AQTAFVSLALFNILRPFPLNLPWVSISSVQASVSLKELRIFLSHEELPDSIER-RPVKO	637
DB	549	PAIVFPSLSLNFILNSAIYSVPSMINTIITSVSMERLKFLLSDEIDDSFIEDPSAD	608
QY	638	GGGTNSITVRNATFTW-----ARSDPPTLNGIT-----FSIPEGALV	674
DB	609	ERALPALTEMNITFLWKSKEVLTSQSGDNLRDDESIIGSSQIALKNIDHFEAKRGDLV	668
QY	675	AVGQVCGCKSSLLSALLAEMDKVEGH-----VAIKGSVAVVPOQAMIQNDLSRENI	726
DB	669	CVVGRVAGKSTFLKAILGQLPCMSGSRDSIPPKLIIRSSSVAYCSQESWIMNASVRENI	728
QY	727	LFGQLEBPYRVIQACALLPDLLEILPSGDRTEIGKGVNLSGQKORVSLARAVYNA	786
DB	729	LFGHKFDQYDYLTIKACQLLPDLKILPDGDETUVGKGLSLSGQKARLSARAVYNA	788
QY	787	DIYLFDDPLSAVDAAHVGKHIPENV-IGPKGMLKNKTRILVTHMSYLPQVDVVIIVMSGGK	845
DB	789	DIYLLDDLSAVDAEVSNIIEYVLIGKTALLKNKTIITNTVSIKHSQMIYALENCE	848
QY	846	ISEMGSYQELLAR---DGAFAEFLRTYAS-----TEQODAEENGVTGVSQPGKEAKQME	897
DB	849	IVEQGNVEDVMNRKNNTSKLKLLEEDSPIDNGNESDVQTEHRSSEVDEP-----	900
QY	898	NGMLVDSAGQLOQORQSSSSSYSGDISRHNSHNTAEIQKABAKKEETWKL-----	947
DB	901	-----LQKVTSETEDEVVT-----ESELIKANSRRASLATLAPRPVGAQL	945
QY	948	-----NEADKAQTQGVKLSYMDYMKALGLFISLFIPLM-CNHVSALASNYWLSLW	999
DB	946	DSVKTAQKAETEVGRVKTILYAIKACGV-LGVVLFLLFMLTRVDFDLAENFWLKYW	1004
QY	1000	TDDPIVNGTQEHKTVRLSVYCALGISQGIAPVGYSMAYSI---GGILASRCLHVDLLHSI	1056
DB	1005	SESNEKNGSNERVWVFGVYSLIGVAS--AAFNLRISIMLLYCSIRGSKKLHESMAKSV	1062
QY	1057	LRSPMSFERPSPGNLNRFSKELDTVDSPITPEVIKMFMGSLFNIVGACIVILLATPIAA	1116
DB	1063	IRSPMTFTETTPVGRHIIIRFSDMDAVDSNLQYIFSFYFKSLITLVTLVILVGNWPL	1122
QY	1117	IIIPPLGIYFPVQRYVASSRQLRLESVRSRSPVSHFNFTLLGVSVIRAPBEOERFIH	1176
DB	1123	VENMELVVIYIYQTFYIVLSRELKRLISISYSPIMSLMSESLNGYSIIDAYDHPERFIY	1182
QY	1177	QSDLKVDENQKAYYPSIVANRWLAVRLSCVGNICVLPAALPAVI---SRHLSAGLVLGUS	1233
DB	1183	LNVEKIQVNVDFVFNFRNTRNLVSRLOITIGATITLATALATATWNTKQLSSGWGLL	1242
QY	1234	VSYSLOVTTYLNLVMSSEMETNIVAVERLKEYSETEKAPWQIQETAPSSWPQVGRV	1293
DB	1243	MSYSEVTSGLTWIVRTVTIETNIVSERIVEYCELPPEAQSIINPEKRPDENWPSKGI	1302
QY	1294	EFNYCLRYREDLDFVLRHINVTINGEKVIGVGTGACKSILTLGLFRINSEAGEIIL	1353
DB	1303	EFKYSTKYRENLDPLVANNINVKIBPCEKVGIVGRTGACKSTLSLALFLEPTEGKIIL	1362
QY	1354	DGINIAKIGLHDLRPFKTIIPQDPVLFSGSLRMLNLPFQSYSDSEEWTSLELAHLKDFVS	1413

Db 1363 DGIDISDIGLFLRSHLAIIPQAFEGTVKTNLDPFNRYSEDELKRAVEQAHLKPHLE 1422  
QY 1414 AL-----PDKLDEHCEAGGENLSVGQRQLVCLARALLRKTILVILDEA 1456  
Db 1423 KMLHSPRGDDSNEDGNVNDILDVKINENGNSLSVGQRQLCLARALLNRSKILVILDEA 1482  
QY 1457 TAAVDLETDDLIQSTIRTOTPEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLIQ 1516  
Db 1483 TASVDMETDKIIQDTIRREFKORTILTIARHRTDVLDSKIIIVLDQGSVREFDPSFKLLS 1542  
QY 1517 QR-GLFYSMAXDAG 1529  
Db 1543 DKTSIFYSLCEKGG 1556

Search completed: March 18, 2005, 11:05:44  
Job time : 62.7349 secB



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 10:57:20 ; Search time 206.058 Seconds  
(without alignments)  
4838.524 Million cell updates/sec

Title: US-10-665-283-8  
Perfect score: 10016  
Sequence: 1 MALRGFCSADGSDPLNDWNV.....RSVAVAKPKFISPDLSL 1947

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	78.5	1531	1 MRPI_HUMAN	P33527 homo sapien
2	7769	77.6	1515	2 Q9UQ99	Q9UQ99 homo sapien
3	7695	76.8	1531	2 Q864R9	Q864R9 macaca fasc
4	7689	76.8	1531	2 Q864S0	Q864S0 macaca fasc
5	7444	74.3	1459	2 Q9UQ97	Q9UQ97 homo sapien
6	7419.5	74.1	1456	2 Q9UQ90	Q9UQ90 homo sapien
7	7272	72.6	1531	2 Q6UR05	Q6UR05 canis fami
8	7230.5	72.2	1530	2 Q8HXQ5	Q8HXQ5 bos taurus
9	7104.5	70.9	1400	2 Q9UQ98	Q9UQ98 homo sapien
10	7002.5	69.9	1528	1 MRPI_MOUSE	O35379 mus musculu
11	6932.5	69.2	1532	2 Q810E4	Q810E4 rattus norv
12	6932.5	69.2	1532	2 Q8CG09	Q8CG09 rattus norv
13	6892	68.8	1523	2 Q810G9	Q810G9 rattus norv
14	6075	60.7	1215	2 Q68CP7	Q68CP7 homo sapien
15	4522	45.1	1519	2 Q80ZK8	Q80ZK8 mus musculu
16	4491.5	44.8	1527	1 MRPI_HUMAN	O15438 homo sapien
17	4422	44.1	1514	2 Q96QA9	Q96QA9 homo sapien
18	4396	43.9	1522	1 MRP3_RAT	Q88563 rattus norv
19	3952.5	39.5	1592	2 Q7Q1D7	Q7Q1D7 anophelis g
20	3885.5	38.8	1548	2 Q7KTC3	Q7KTC3 drosophila
21	3880.5	38.7	1548	2 Q7KTC1	Q7KTC1 drosophila
22	3879.5	38.7	1548	2 Q7KTB7	Q7KTB7 drosophila
23	3874.5	38.7	1548	2 Q7KTC2	Q7KTC2 drosophila
24	3860.5	38.5	1548	2 Q917N0	Q917N0 drosophila
25	3857	38.5	1549	2 Q7KTC0	Q7KTC0 drosophila
26	3845.5	38.4	1548	2 Q8T9C5	Q8T9C5 drosophila
27	3838.5	38.3	1548	2 Q7KTB8	Q7KTB8 drosophila
28	3836.5	38.3	1548	2 Q7KTB9	Q7KTB9 drosophila
29	3818.5	38.1	1548	2 Q7KTD0	Q7KTD0 drosophila
30	3813.5	38.1	1548	2 Q7KTC8	Q7KTC8 drosophila
31	3812.5	38.1	1548	2 Q7KTC4	Q7KTC4 drosophila

RESULT 1

ID	MRPI_HUMAN	STANDARD;	PRT;	1531 AA.
AC	P33527; O14819; P78419;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Multidrug resistance-associated protein 1 (ATP-binding cassette, sub-family C, member 1).			
DE	Name=ABCC1; Synonyms=MRP, MRPI;			
GN	Homo sapiens (Human).			
OS	Sukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9308080; PubMed=1360704;			
RA	Cole S.P.C., Bhargava J.G., Gerlach J.H., Mackie J.E., Grant C.E., Almqvist K.C., Stewart A.J., Kurz E.U., Duncan A.M.V., Deeley R.G.;			
RT	"Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line."			
RL	Science 258:1650-1654(1992).			
RN	[2]			
RP	REVISIONS.			
RX	MEDLINE=93262415; PubMed=8098549;			
RA	Cole S.P.C., Deeley R.G.;			
RT	"Multidrug resistance-associated protein: sequence correction."			
RL	Science 260:879-879(1993).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96251691; PubMed=8649356;			
RA	Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C., Deeley R.G.;			
RT	"Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter."			
RL	Mol. Pharmacol. 49:962-971(1996).			
RN	[4]			
RP	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.			
RX	MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;			
RA	Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;			
RT	"Analysis of the intron-exon organization of the human multidrug-resistance protein gene (MRP) and alternative splicing of its mRNA."			
RL	Genomics 45:368-378(1997).			
RN	[5]			
RP	SEQUENCE OF 1131-1531 FROM N.A.			
RX	MEDLINE=99425270; PubMed=10493829; DOI=10.1006/geno.1999.5927;			
RA	Lofthus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R., Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.B., Harris P.C., Venter J.C., Adams M.D.;			
RT	"Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q."			
RL	Genomics 60:295-308(1999).			
RN	[6]			

32	3807.5	38.0	1548	2	Q7KTC9	O7ktc9 drosophila
33	3793.5	37.9	1548	2	Q7KTC7	O7ktc7 drosophila
34	3790	37.8	1549	2	Q9VK56	O9vk56 drosophila
35	3771.5	37.7	1548	2	Q7KTC5	O7ktc5 drosophila
36	3769.5	37.6	1548	2	Q7KTC6	O7ktc6 drosophila
37	3741	37.4	1564	2	Q8QG98	Q8qg98 raja erinac
38	3738.5	37.3	1567	2	O6PH26	O6ph26 brachydanio
39	3655	36.5	822	2	O9JHS0	O9jhs0 rattus norv
40	3653.5	36.5	1544	2	O6PSM3	O6psm3 canis fami
41	3641.5	36.4	1544	2	O9SM36	O9sm36 canis fami
42	3625.5	36.2	796	2	Q8C7V6	Q8c7v6 mus musculu
43	3595	35.9	1545	1	MRP2_HUMAN	Q92887 homo sapien
44	3594	35.9	1544	2	Q95L75	Q95l75 macaca mula
45	3577	35.7	1564	1	MRP2_RABIT	Q28689 o canalicul

ALIGNMENTS

TOPOLOGY, AND CARBOHYDRATE-LINKAGE SITES.  
 RX MEDLINE=97442425; PubMed=9295302; DOI=10.1074/jbc.272.38.23623;  
 RA Hipfner D.R., Almqvist K.C., Leslie E.M., Gerlach J.H., Grant C.B.,  
 RA Deeley R.G., Cole S.P.C.;  
 RT "Membrane topology of the multidrug resistance protein (MRP). A study  
 RT of glycosylation-site mutants reveals an extracytosolic NH2  
 RT terminus.";  
 RL J. Biol. Chem. 272:23623-23630(1997).  
 RN [7]

TOPOLOGY.  
 RP MEDLINE=97476249; PubMed=9334225; DOI=10.1074/jbc.272.42.26479;  
 RX Kast C., Gros P.;  
 RA "Topology mapping of the amino-terminal half of multidrug resistance-  
 RT associated protein by epitope insertion and immunofluorescence.";  
 RL J. Biol. Chem. 272:26479-26487(1997).  
 RN [8]

TOPOLOGY.  
 RP MEDLINE=98153110; PubMed=9485377; DOI=10.1021/bi972332v;  
 RX Kast C., Gros P.;  
 RA "Epitope insertion favors a six transmembrane domain model for the  
 RT carboxy-terminal portion of the multidrug resistance-associated  
 RT protein.";  
 RL Biochemistry 37:2305-2313(1998).  
 RN [9]

MUTAGENESIS OF ASP-792; ASP-793; LYS-1333 AND 1454-ASP-GLU-1455.  
 RP MEDLINE=21362977; PubMed=11469806; DOI=10.1006/abbi.2001.2441;  
 RX Cui L., Hou Y.-X., Riordan J.R., Chang X.-B.;  
 RA "Mutations of the Walker B motif in the first nucleotide binding  
 RT domain of multidrug resistance protein MRP1 prevent conformational  
 RT maturation.";  
 RL Arch. Biochem. Biophys. 392:153-161(2001).  
 RN [10]

MUTAGENESIS OF TRP-1246.  
 RP MEDLINE=21238219; PubMed=11728867; DOI=10.1074/jbc.M011246200;  
 RX Ito K., Olsen S.L., Qiu W., Deeley R.G., Cole S.P.C.;  
 RA "Mutation of a single conserved tryptophan in multidrug resistance  
 RT protein 1 (MRP1/ABCC1) results in loss of drug resistance and  
 RT selective loss of organic anion transport.";  
 RL J. Biol. Chem. 276:15616-15624(2001).  
 RN [11]

VARIANTS GLN-633 AND VAL-671.  
 RP MEDLINE=20283940; PubMed=10835642; DOI=10.1038/76102;  
 RX Le Saux O., Urban Z., Tschuch C., Csizsar K., Sacchelli B.,  
 RA Quagliaro D., Pasquali-Ronchetti I., Pope F.M., Richards A., Terry S.,  
 RA Bercovich L., de Paape A., Boyd C.D.;  
 RT "Mutations in a gene encoding an ABC transporter cause pseudoxanthoma  
 RT elasticum.";  
 RL Nat. Genet. 25:223-227(2000).  
 RN [12]

VARIANT VAL-671.  
 RP MEDLINE=20283940; PubMed=10835642; DOI=10.1073/pnas.100041297;  
 RX Ringfeil F., Lebowitz M.G., Cristiano A.M., Uitto J.;  
 RA "Pseudoxanthoma elasticum: mutations in the MRP6 gene encoding a  
 RT transmembrane ATP-binding cassette (ABC) transporter.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:6001-6006(2000).  
 RN [13]

VARIANT SER-433, AND CHARACTERIZATION OF VARIANT VAL-671.  
 RP MEDLINE=21578494; PubMed=11721885;  
 RX Conrad S., Kauffmann H.-M., Ito K., Deeley R.G., Cole S.P.C.,  
 RA Schrenk D.;  
 RT "Identification of human multidrug resistance protein 1 (MRP1)  
 RT mutations and characterization of a G671V substitution.";  
 RL J. Hum. Genet. 46:656-663(2001).  
 RN [14]

VARIANTS MET-117 AND LEU-1512.  
 RP MEDLINE=20579883; PubMed=11139250;  
 RX DOI=10.1002/1098-1004(2001)17:1<74::AID-HUMU14-3.0.CO;2-F;  
 RA Perdu J., Germain D.P.;  
 RT "Identification of novel polymorphisms in the pM5 and MRP1 (ABCC1)  
 RT genes at locus 16p13.1 and exclusion of both genes as responsible for  
 RT pseudoxanthoma elasticum.";  
 RL Hum. Mutat. 17:74-75(2001).  
 RN [15]

VARIANTS SER-433; ILE-73; GLN-723 AND GLN-1058.  
 RP MEDLINE=21163948; PubMed=11266082;  
 RX Ito S., Ieiri I., Tanabe M., Suzuki A., Higuchi S., Otsubo K.;  
 RA "Polymorphism of the ABC transporter genes, MDR1, MRP1 and MRP2/cMOAT,  
 RT in healthy Japanese subjects.";  
 RL Pharmacogenetics 11:175-184(2001).  
 RN [16]

FUNCTION: May participate directly in the active transport of  
 drugs into subcellular organelles or influence drug distribution  
 indirectly.  
 CC SUBCELLULAR LOCATION: Integral membrane protein.  
 CC ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing: Named isoforms=8;  
 CC Comment-Additional isoforms seem to exist. Experimental  
 CC confirmation may be lacking for some isoforms;  
 CC Name=Alleles;  
 CC IsoId=P33527-1; Sequence=Displayed;  
 CC Name=Delexon-17;  
 CC IsoId=P33527-2; Sequence=VSP\_000037;  
 CC Name=Delexon-18;  
 CC IsoId=P33527-3; Sequence=VSP\_000038;  
 CC Name=Delexon-30;  
 CC IsoId=P33527-4; Sequence=VSP\_000039;  
 CC Name=Delexon-17-18;  
 CC IsoId=P33527-5; Sequence=VSP\_000037, VSP\_000038;  
 CC Name=Delexon-17-30;  
 CC IsoId=P33527-6; Sequence=VSP\_000037, VSP\_000039;  
 CC Name=Delexon-18-30;  
 CC IsoId=P33527-7; Sequence=VSP\_000038, VSP\_000039;  
 CC Name=Delexon-17-18-30;  
 CC IsoId=P33527-8; Sequence=VSP\_000037, VSP\_000038, VSP\_000039;  
 CC TISSUE SPECIFICITY: Lung, testis and peripheral blood mononuclear  
 CC cells.  
 CC SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.  
 CC DATABASE: NAME=Atlas Genet. Cytoget. Oncol. Haematol.;  
 CC WWW=HTTP://www.infobiogen.fr/services/Chromancer/Genes/MRP106.html".  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L05628; AAB46616.1; --  
 CC EMBL; AF022853; AAB83983.1; JOINED.  
 CC EMBL; AF022824; AAB83983.1; JOINED.  
 CC EMBL; AF022825; AAB83983.1; JOINED.  
 CC EMBL; AF022826; AAB83983.1; JOINED.  
 CC EMBL; AF022827; AAB83983.1; JOINED.  
 CC EMBL; AF022828; AAB83983.1; JOINED.  
 CC EMBL; AF022829; AAB83983.1; JOINED.  
 CC EMBL; AF022830; AAB83983.1; JOINED.  
 CC EMBL; AF022831; AAB83983.1; JOINED.  
 CC EMBL; AF022832; AAB83983.1; JOINED.  
 CC EMBL; AF022833; AAB83983.1; JOINED.  
 CC EMBL; AF022834; AAB83983.1; JOINED.  
 CC EMBL; AF022835; AAB83983.1; JOINED.  
 CC EMBL; AF022836; AAB83983.1; JOINED.  
 CC EMBL; AF022837; AAB83983.1; JOINED.  
 CC EMBL; AF022838; AAB83983.1; JOINED.  
 CC EMBL; AF022839; AAB83983.1; JOINED.  
 CC EMBL; AF022840; AAB83983.1; JOINED.  
 CC EMBL; AF022841; AAB83983.1; JOINED.  
 CC EMBL; AF022842; AAB83983.1; JOINED.  
 CC EMBL; AF022843; AAB83983.1; JOINED.  
 CC EMBL; AF022844; AAB83983.1; JOINED.  
 CC EMBL; AF022845; AAB83983.1; JOINED.  
 CC EMBL; AF022846; AAB83983.1; JOINED.  
 CC EMBL; AF022847; AAB83983.1; JOINED.  
 CC EMBL; AF022848; AAB83983.1; JOINED.  
 CC EMBL; AF022849; AAB83983.1; JOINED.  
 CC EMBL; AF022850; AAB83983.1; JOINED.

961	SVYDWMKAIGLFI	SPLSIFLFCMNVHSALASNYWLSLTD	DDPIVNGTQHTTKVRLSVYG	102
1021	ALGISOGIAVFGYSMAVSI	GGIILASRCLHVDLLHLSILRS	PMSFPERTPSGNLVNRFSEKEL	1080
1021	ALGISOGIAVFGYSMAVSI	GGIILASRCLHVDLLHLSILRS	PMSFPERTPSGNLVNRFSEKEL	1080
1081	DTVDSMIPEVIKMFMSGLFN	VIACIVILATPIAAIIIPPLGLIYF	VFQRFYVASSRQL	1140
1081	DTVDSMIPEVIKMFMSGLFN	VIACIVILATPIAAIIIPPLGLIYF	VFQRFYVASSRQL	1140
1141	KRLSVSRSPVYSHFNETHL	IGSVIRAFEFQERFIHOSDLKVDEN	KAYYPSIVANRWLA	1200
1141	KRLSVSRSPVYSHFNETHL	IGSVIRAFEFQERFIHOSDLKVDEN	KAYYPSIVANRWLA	1200
1201	VRLECVGNCIVLFAALFAV	ISRHSLSAGLVGLSVSYSLQVTT	YLNVLVRMSSEMETNIVA	1260
1201	VRLECVGNCIVLFAALFAV	ISRHSLSAGLVGLSVSYSLQVTT	YLNVLVRMSSEMETNIVA	1260
1261	VERLKEYSETEKAPWQIQ	BTAPSSWPQGRVFRNYCLRYRED	LDVLRHINVTINGG	1320
1261	VERLKEYSETEKAPWQIQ	BTAPSSWPQGRVFRNYCLRYRED	LDVLRHINVTINGG	1320
1321	EKGIVGRTGAGKSSLTGL	FRINESAEGEIIIDGINIAKIGLH	DLRFKTIIPQDPVLF	1380
1321	EKGIVGRTGAGKSSLTGL	FRINESAEGEIIIDGINIAKIGLH	DLRFKTIIPQDPVLF	1380
1381	SGSLRWMLDPFSQYSD	BEVWTSLELAHLKDFVSALPKL	DHECAEGGENLSVGQRQLVCL	1440
1381	SGSLRWMLDPFSQYSD	BEVWTSLELAHLKDFVSALPKL	DHECAEGGENLSVGQRQLVCL	1440
1441	ARALLRKTILVLDEATA	AVDLETDLLIQSTIRTPQEDCT	VLTTIAHRLNTIMDTRVIVL	1500
1441	ARALLRKTILVLDEATA	AVDLETDLLIQSTIRTPQEDCT	VLTTIAHRLNTIMDTRVIVL	1500
1501	DKGEIQYAGAPSDLLQ	QRGLFYFSMAKDAGLV	1531	
1501	DKGEIQYAGAPSDLLQ	QRGLFYFSMAKDAGLV	1531	

RESULT 2

Q9UQ99 ID Q9UQ99 PRELIMINARY; PRT; 1515 AA.

AC Q9UQ99;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Multidrug resistance protein (Fragment).

GN Name=MRP;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RN SEQUENCE FROM N.A.

RX MEDLINE=98008927; PubMed=3344662; DOI=10.1006/geno.1997.4950;

RA Grant C.E., Kurz E.U., Cole S.P., Deeley R.G.;

RT "Analysis of the intron-exon organization of the human multidrug-

RT resistance protein gene (MRP) and alternative splicing of its mRNA.";

RL Genomics 45:368-378(1997).

CC -1- SIMILARITY: Belongs to the ABC transporter family.

DR EMBL; AF022827; AAB83979.1; JOINED.

DR EMBL; AF022828; AAB83979.1; JOINED.

DR EMBL; AF022829; AAB83979.1; JOINED.

DR EMBL; AF022830; AAB83979.1; JOINED.

DR EMBL; AF022831; AAB83979.1; JOINED.

DR EMBL; AF022832; AAB83979.1; JOINED.

DR EMBL; AF022833; AAB83979.1; JOINED.

DR EMBL; AF022834; AAB83979.1; JOINED.

DR EMBL; AF022835; AAB83979.1; JOINED.

DR EMBL; AF022836; AAB83979.1; JOINED.

DR EMBL; AF022837; AAB83979.1; JOINED.

DR EMBL; AF022838; AAB83979.1; JOINED.

DR EMBL; AF022839; AAB83979.1; JOINED.

DR EMBL; AF022840; AAB83979.1; JOINED.  
 DR DR EMBL; AF022841; AAB83979.1; JOINED.  
 DR DR EMBL; AF022842; AAB83979.1; JOINED.  
 DR DR EMBL; AF022843; AAB83979.1; JOINED.  
 DR DR EMBL; AF022844; AAB83979.1; JOINED.  
 DR DR EMBL; AF022845; AAB83979.1; JOINED.  
 DR DR EMBL; AF022846; AAB83979.1; JOINED.  
 DR DR EMBL; AF022847; AAB83979.1; JOINED.  
 DR DR EMBL; AF022848; AAB83979.1; JOINED.  
 DR DR EMBL; AF022849; AAB83979.1; JOINED.  
 DR DR EMBL; AF022850; AAB83979.1; JOINED.  
 DR DR EMBL; AF022851; AAB83979.1; JOINED.  
 DR DR EMBL; AF022852; AAB83979.1; JOINED.  
 DR DR EMBL; AF022853; AAB83979.1; JOINED.  
 DR HSSP; P08716; IMT0.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005224; F:ATP binding; IEA.  
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO; GO:000166; F:nucleotide binding; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR011527; ABC membrane 1.  
 DR InterPro; IPR001140; ABC TM transp.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR Pfam; PF00664; ABC membrane; 2.  
 DR Pfam; PF00005; ABC tran; 2.  
 DR ProDom; PD000006; ABC transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAM; TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE; PS00329; ABC\_TMIF; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 KW ATP-binding.  
 FT NON\_TER  
 SQ SEQUENCE 1 169852 MW; DC85592817C439FE CRC64;  
 Query Match 77.6%; Score 7769; DB 2; Length 1515;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 17 DNNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFFLYLSRHRDGYIQMTPLNKTCTA 76  
 DB 1 DNNVTWNTSNPDFTKCFQNTVLVWVPCFYLWACFPFFFLYLSRHRDGYIQMTPLNKTCTA 60  
 QY 77 LGFLLIWVCWADLFYSFWRSGRIFLAPVFLVSPITLLGTTLLATFLIQLERRKGVQSSG 136  
 DB 61 LGFLLIWVCWADLFYSFWRSGRIFLAPVFLVSPITLLGTTLLATFLIQLERRKGVQSSG 120  
 QY 137 IMLTFLWALVCAILRSKIMTALKEDAQVDLPDITFYFYFSLILLQLVLSCFSDRSP 196  
 DB 121 IMLTFLWALVCAILRSKIMTALKEDAQVDLPDITFYFYFSLILLQLVLSCFSDRSP 180  
 QY 197 LFSETHDPNCPSSASFLSRITFWITGLIVRGYRQPLEGSLWSLNKEDTSEQVVPV 256  
 DB 181 LFSETHDPNCPSSASFLSRITFWITGLIVRGYRQPLEGSLWSLNKEDTSEQVVPV 240  
 QY 257 LVKNWKECAKTRQPKVWVSSKDPAPQKSSKVDANEVEALIVKSPQKWNPSLFKV 316  
 DB 241 LVKNWKECAKTRQPKVWVSSKDPAPQKSSKVDANEVEALIVKSPQKWNPSLFKV 300  
 QY 317 LYKTEGPVFLMSFFFKAIHDLMMFSGPQILKLIKFNVDTKAPDQWGYFTVLLFVTA 376  
 DB 301 LYKTEGPVFLMSFFFKAIHDLMMFSGPQILKLIKFNVDTKAPDQWGYFTVLLFVTA 360  
 QY 377 QTLVLHQYFHCIFVSGMEIKTAVTGAVYRKALVITNSARKSSTVGEIVNLMSVDQRFMD 436

DB 361 QTLVLHQYFHCIFVSGMEIKTAVTGAVYRKALVITNSARKSSTVGEIVNLMSVDQRFMD 420  
 QY 437 LATYINMIWSAPLQVILALYLLNLGSPSVLAGVAVMVLMPVNVANAMKTKTYQVAHMK 496  
 DB 421 LATYINMIWSAPLQVILALYLLNLGSPSVLAGVAVMVLMPVNVANAMKTKTYQVAHMK 480  
 QY 497 SKDNRIKLMMNELNGIKVLKYAWELAPDKVLAIROBELKVLKKSAYLSAVGFTTWCT 556  
 DB 481 SKDNRIKLMMNELNGIKVLKYAWELAPDKVLAIROBELKVLKKSAYLSAVGFTTWCT 540  
 QY 557 PFLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRPFLNIPMWISSIVQASVSLKRL 616  
 DB 541 PFLVALCTFAVYVITDENNILDQAOTAFVSLALFNILRPFLNIPMWISSIVQASVSLKRL 600  
 QY 617 RIFLSHEBLEPDSITERRPVKGGGTNSITVRNATFTWARSDDPTINGITFTFPEGALVAV 676  
 DB 601 RIFLSHEBLEPDSITERRPVKGGGTNSITVRNATFTWARSDDPTINGITFTFPEGALVAV 660  
 QY 677 VGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWIQNDLSRENILFGCOLREPY 736  
 DB 661 VGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWIQNDLSRENILFGCOLREPY 720  
 QY 737 YRSVIQACALLPDLLEILPSPGDRTEIGEGVNLSGQKQVSLARAVYSNADIYLFDDPLS 796  
 DB 721 YRSVIQACALLPDLLEILPSPGDRTEIGEGVNLSGQKQVSLARAVYSNADIYLFDDPLS 780  
 QY 797 AVDAHVGKHFENVIGPKGMLKNKTRILVTHSMYLPQVDVIIVMGSGKISEMGSYQELL 856  
 DB 781 AVDAHVGKHFENVIGPKGMLKNKTRILVTHSMYLPQVDVIIVMGSGKISEMGSYQELL 840  
 QY 857 ARDGAFAEFLRTYASTEOQDAENGVTGSGPCKEAKQEMGNMLVTDGAKQLQROLSS 916  
 DB 841 ARDGAFAEFLRTYASTEOQDAENGVTGSGPCKEAKQEMGNMLVTDGAKQLQROLSS 900  
 QY 917 SSSYSGDISRHHNSTABEQKAEAKKEETWKLMEADKATQGVKLSVYMDYKMAIGLFI 976  
 DB 901 SSSYSGDISRHHNSTABEQKAEAKKEETWKLMEADKATQGVKLSVYMDYKMAIGLFI 960  
 QY 977 LSIFLFCNHSVASALSNYLSLWTDPIVNGTQHTKVRLSVYGALGISQIAVFGYSMA 1036  
 DB 961 LSIFLFCNHSVASALSNYLSLWTDPIVNGTQHTKVRLSVYGALGISQIAVFGYSMA 1020  
 QY 1037 VSIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKELTDVDSMIPKMFMG 1096  
 DB 1021 VSIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFSKELTDVDSMIPKMFMG 1080  
 QY 1097 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFVYVASSRQLKRLESVSRSPVSHFN 1156  
 DB 1081 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFVYVASSRQLKRLESVSRSPVSHFN 1140  
 QY 1157 ETLGVSIVIRAFEOERFIHQSDLKVDENOKAYYPSIVANRWLAVRLECVGNCIVLFAAL 1216  
 DB 1141 ETLGVSIVIRAFEOERFIHQSDLKVDENOKAYYPSIVANRWLAVRLECVGNCIVLFAAL 1200  
 QY 1217 FAVISRHSLSAGLVGLSVSYSLQVTTYINLVRMSSEMETNIVAVERLKEYSETEKAPW 1276  
 DB 1201 FAVISRHSLSAGLVGLSVSYSLQVTTYINLVRMSSEMETNIVAVERLKEYSETEKAPW 1260  
 QY 1277 QIQTAPSSWPQVGRVEFRNYCLRYREDLDVLRHINVTINGEKVIGIVGRTGAGKSSL 1336  
 DB 1261 QIQTAPSSWPQVGRVEFRNYCLRYREDLDVLRHINVTINGEKVIGIVGRTGAGKSSL 1320  
 QY 1337 TLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLFGSLRNWLPFSSQYSD 1396  
 DB 1321 TLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLFGSLRNWLPFSSQYSD 1380  
 QY 1397 EEWVTSLELAHLKDFVSALPDKLDHECAEGENLSVQROLVCLARALKRKTILVDEA 1456  
 DB 1381 EEWVTSLELAHLKDFVSALPDKLDHECAEGENLSVQROLVCLARALKRKTILVDEA 1440  
 QY 1457 TAAVDLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1516  
 DB 1441 TAAVDLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1500

```

QY 1517 ORGLFYSMAKDAGLV 1531
DB 1501 ORGLFYSMAKDAGLV 1515

RESULT 3
Q864R9
ID Q864R9 PRELIMINARY; PRT; 1531 AA.
AC Q864R9;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Multidrug resistance protein 1B.
GN Name=MRP1;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22544876; PubMed=12657726;
RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RA Dantzig A.H., Perry W.L.;
RT "Cloning and functional characterization of the multidrug resistance-
RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
RL Mol. Cancer Ther. 2:307-316(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Perry W.L. III, Godinot N.;
RA Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY146673; AAN65349.1; -
DR HSP; P08716; 1MTO.
DR GO; GO:0015021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TWIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50993; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding.
SQ SEQUENCE 1531 AA; 171658 MW; 1AE788BFD9EF459 CRC64;

Query Match 76.8%; Score 7695; DB 2; Length 1531;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 1493; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLWDMNVNTSNPDTKCFQNTLVWVPCFYLWACFPFYLYLSRH 60
DB 1 MALRGFCSADGSDPLWDMNVNTSNPDTKCFQNTLVWVPCFYLWACFPFYLYLSRH 60
QY 61 DRGYIQMTPLNKTALGFLWVWADLFYFWSRSGIFLAPVFLVSPITLGLITLLA 120
DB 61 DRGYIQMTPLNKTALGFLWVWADLFYFWSRSGIFLAPVFLVSPITLGLITLLA 120
QY 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDQVVDLFRDITFYVPS 180
DB 121 TFLIQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDQVVDLFRDITFYVPS 180

```

QY 1261 VERLKEYSETEAPWQIOETAPPSWPOVGRVFRNYCLRYREDLDFVLRHINTVINGG 1320  
 Db 1261 VERLKEYSETEAPWQIOETAPPSWPOVGRVFRNYCLRYREDLDFVLRHINTVINGG 1320  
 QY 1321 EKVGIVGRTGACKSSLTGLPRINSAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 Db 1321 EKVGIVGRTGACKSSLTGLPRINSAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGLSRNLDPPFOYSDDEEVTSLAHLKDFYSALPKLDHCAEGGENLSVGQRLVCL 1440  
 Db 1381 SGLSRNLDPPFOYSDDEEVTSLAHLKDFYSALPKLDHCAEGGENLSVGQRLVCL 1440  
 QY 1441 ARALLRKTILVLDATAVLETDLQIOTRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLRKTILVLDATAVLETDLQIOTRTQFEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DXGEIQEYGAPSDLLQORGLFYSMADAGLV 1531  
 Db 1501 DXGEIQEYGAPSDLLQORGLFYNMARDAGLV 1531

RESULT 4

Q864SO PRELIMINARY; PRT; 1531 AA.  
 AC Q864SO;  
 DT 01-JUN-2003 (TremBLrel. 24, Created)  
 DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)  
 DT 01-WAR-2004 (TremBLrel. 26, Last annotation update)  
 DE Multidrug resistance protein 1A.  
 GN Name=MRP1;  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
 OC Cercopitheciinae; Macaca.  
 OX NCBI\_taxid=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22544876; PubMed=12657726;  
 RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,  
 RA Dantzig A.H., Perry W.L.;  
 RT "Cloning and functional characterization of the multidrug resistance-  
 RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.";  
 RL Mol. Cancer Ther. 2:307-316(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Perry W.L., Ill, Godinot N.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AY146672; AAN65348.1; -;  
 DR HSSP; P08716; 1MT0.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR011527; ABC membrane 1.  
 DR InterPro; IPR001140; ABC TM transp.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR005292; MRP assoc.  
 DR InterPro; IPR000719; Prot.Kinase.  
 DR Pfam; PF00664; ABC membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR Prodom; PD00006; ABC transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAMs; TIGR00957; MRP assoc pro; 1.  
 DR PROSITE; PS00929; ABC TMIF; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.

DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 KW ATP-binding.  
 SQ SEQUENCE 1531 AA; 171700 MW; 16871288CC2D2B89 CRC64;  
 Query Match 76.8%; Score 7689; DB 2; Length 1531;  
 Best Local Similarity 97.6%; Pred. No. 0;  
 Matches 1494; Conservative 18; Mismatches 19; Indels 0; Gaps 0;  
 QY 1 MALRGFCADSGDPLDWNVTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFFLYLSRH 60  
 Db 1 MALRGFCADSGDPLDWNVTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFFLYLSRH 60  
 QY 61 DRGYIQTMLNKTALGFLIWCWADLFYSFERSRGIFLAPVFLVSPITLLGTTLLA 120  
 Db 61 DRGYIQTMLNKTALGFLIWCWADLFYSFERSRGIFLAPVFLVSPITLLGTTLLA 120  
 QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 Db 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 QY 181 LLLIQLVLSCESDRSPLEFSETHDNPCESSASFLSRITFWWITGLIVRGYRPLEGSD 240  
 Db 181 LLLIQLVLSCESDRSPLEFSETHDNPCESSASFLSRITFWWITGLIVRGYRPLEGSD 240  
 QY 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVWVSSKDPAPQPKSSKVDANEEVEAL 300  
 Db 241 LWSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVWVSSKDPAPQPKSSKVDANEEVEAL 300  
 QY 301 IVKSPQKWNPSLFKVLKTFPGPYFLMSFPFKAHDLMPFSGPOILKLLIKFVNDTKAPD 360  
 Db 301 IVKSPQKWNPSLFKVLKTFPGPYFLMSFPFKAHDLMPFSGPOILKLLIKFVNDTKAPD 360  
 QY 361 WQGYFYTVLLFVTAACLOTVLHOFVHICFVSGMEIKTAVIGAVYRKALVITNSARKSTV 420  
 Db 361 WQGYFYTVLLFVTAACLOTVLHOFVHICFVSGMEIKTAVIGAVYRKALVITNSARKSTV 420  
 QY 421 GEIVNLSVDAQREFMDLATYINMIWSAPLOVILALYLLNMLNLSVLAGVAVMLVMPVN 480  
 Db 421 GEIVNLSVDAQREFMDLATYINMIWSAPLOVILALYLLNMLNLSVLAGVAVMLVMPVN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLK 540  
 QY 541 KSAVLSAVGTFTWCTPFLVALCTFVAVVTIDENNILDADOTAFVLSLAFNLIRPLNLIP 600  
 Db 541 KSAVLSAVGTFTWCTPFLVALCTFVAVVTIDENNILDADOTAFVLSLAFNLIRPLNLIP 600  
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSDDPT 660  
 QY 661 LINGITFPIREGALVAVVGVQCGCKSSLLSALLAEMDKVEGHVAIKGSVAVYVQAWIOND 720  
 Db 661 LINGITFPIREGALVAVVGVQCGCKSSLLSALLAEMDKVEGHVAIKGSVAVYVQAWIOND 720  
 QY 721 SLRNLIFGQLEPEPYRVSIVIOACALLPDLIELPSGDRTEIGEGVNLSGQKORVSLAR 780  
 Db 721 SLRNLIFGQLEPEPYRVSIVIOACALLPDLIELPSGDRTEIGEGVNLSGQKORVSLAR 780  
 QY 781 AVYSNADIYLFDDPLSAVDAHVGKHFENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIIV 840  
 Db 781 AVYSNADIYLFDDPLSAVDAHVGKHFENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIIV 840  
 QY 841 MSGGKISEMGSYQELLARDGAFELRTYASTQEDAEENGVTGVSGRKEAKOMENGM 900  
 Db 841 MSGGKISEMGSYQELLARDGAFELRTYASTQEDAEENGVTGVSGRKEAKOMENGM 900  
 QY 901 LVTDGAKQIQORQLSSSSSYSGDISRHHNSTAEIQAKAEKKEETWKLMEADKATQGVKL 960  
 Db 901 LVTDGAKQIQORQLSSSSSYSGDISRHHNSTAEIQAKAEKKEETWKLMEADKATQGVKL 960  
 QY 961 SVYWDYMKAIGLFISLISIFLFCNHNVSALASNYWLSLTDTPDPIVNGTQEHKVRLLSYG 1020

Db 961 SVYDYMKAIGLSEISLPLFTCNHVAALASNYLSLWTDITVNGTQHTKVRSLVYG 1020  
 QY 1021 ALGISOGIAVFGYSMAVSGIILASRCLHVDLHLSILRSPMSFFERTPSCNLVNRFSKEL 1080  
 Db 1021 ALGISOGIAVFGYSMAVSGIILASRCLHVDLHLSILRSPMSFFERTPSCNLVNRFSKEL 1080  
 QY 1081 DTVDSDIMEPIKMFMSGLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 Db 1081 DTVDSDIMEPIKMFMSGLFNVIACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQL 1140  
 QY 1141 KRLESVRSRVSYPHFNETLLGVSIVIRAFERQERFIHQSDLKVDENKAYPSPVANRWLA 1200  
 Db 1141 KRLESVRSRVSYPHFNETLLGVSIVIRAFERQERFIHQSDLKVDENKAYPSPVANRWLA 1200  
 QY 1201 VRLECGVNCIVLFAALFAVISRSLSAGLVGLSVSYSLQVTTVNLNLRMSSEMETNIVA 1260  
 Db 1201 VRLECGVNCIVLFAALFAVISRSLSAGLVGLSVSYSLQVTTVNLNLRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKAPQIQETAPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVINGG 1320  
 Db 1261 VERLKEYSETEKAPQIQETAPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVINGG 1320  
 QY 1321 EKVGIVGRTGAGKSSITLGLFRINESAEGEIIIDGINIAKIGHDLRFKTIIPQDPVLV 1380  
 Db 1321 EKVGIVGRTGAGKSSITLGLFRINESAEGEIIIDGINIAKIGHDLRFKTIIPQDPVLV 1380  
 QY 1381 SGLRNLNLPFSSYDEEVTSLLEHLKDFVSALPKLDHCEAEGENLSVQROLVCL 1440  
 Db 1381 SGLRNLNLPFSSYDEEVTSLLEHLKDFVSALPKLDHCEAEGENLSVQROLVCL 1440  
 QY 1441 ARALLKTKILVDEATAVLDLTDLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 Db 1441 ARALLKTKILVDEATAVLDLTDLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIYEGAPSDLLQOGLFYFMAKDAGIV 1531  
 Db 1501 DKGEIYEGAPSDLLQOGLFYFMAKDAGIV 1531

RESULT 5  
 Q9UQ97 PRELIMINARY; PRT; 1459 AA.  
 AC Q9UQ97;  
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
 DE Multidrug resistance protein (fragment).  
 GN Name=MRP;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;  
 RA Grant C.E., Kurz E.U., Cole S.P., Deeley R.G.;  
 RT "Analysis of the intron-exon organization of the human multidrug-  
 RT resistance protein gene (MRP) and alternative splicing of its mRNA";  
 RL Genomics 45:368-378 (1997).  
 CC -!- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AF022824; AAB83980.1; JOINED.  
 DR EMBL; AF022825; AAB83980.1; JOINED.  
 DR EMBL; AF022826; AAB83980.1; JOINED.  
 DR EMBL; AF022827; AAB83980.1; JOINED.  
 DR EMBL; AF022828; AAB83980.1; JOINED.  
 DR EMBL; AF022829; AAB83980.1; JOINED.  
 DR EMBL; AF022830; AAB83980.1; JOINED.  
 DR EMBL; AF022831; AAB83980.1; JOINED.  
 DR EMBL; AF022832; AAB83980.1; JOINED.  
 DR EMBL; AF022833; AAB83980.1; JOINED.  
 DR EMBL; AF022834; AAB83980.1; JOINED.  
 DR EMBL; AF022835; AAB83980.1; JOINED.

DR EMBL; AF022836; AAB83980.1; JOINED.  
 DR EMBL; AF022837; AAB83980.1; JOINED.  
 DR EMBL; AF022838; AAB83980.1; JOINED.  
 DR EMBL; AF022839; AAB83980.1; JOINED.  
 DR EMBL; AF022841; AAB83980.1; JOINED.  
 DR EMBL; AF022842; AAB83980.1; JOINED.  
 DR EMBL; AF022843; AAB83980.1; JOINED.  
 DR EMBL; AF022844; AAB83980.1; JOINED.  
 DR EMBL; AF022845; AAB83980.1; JOINED.  
 DR EMBL; AF022846; AAB83980.1; JOINED.  
 DR EMBL; AF022847; AAB83980.1; JOINED.  
 DR EMBL; AF022848; AAB83980.1; JOINED.  
 DR EMBL; AF022849; AAB83980.1; JOINED.  
 DR EMBL; AF022850; AAB83980.1; JOINED.  
 DR EMBL; AF022851; AAB83980.1; JOINED.  
 DR EMBL; AF022852; AAB83980.1; JOINED.  
 DR EMBL; AF022853; AAB83980.1; JOINED.  
 DR HSSP; P08716; IMT0.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m...; IEA.  
 DR GO; GO:000166; F:nucleotide binding; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR011527; ABC membrane 1.  
 DR InterPro; IPR001140; ABC TM transp.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR Pfam; PF00664; ABC membrane; 2.  
 DR Pfam; PF00005; ABC tran; 1.  
 DR ProDom; PD000006; ABC transporter; 1.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAMs; TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE; PS00929; ABC TMIF; 2.  
 DR PROSITE; PS00211; ABC TRANSPORTER 1; UNKNOWN 1.  
 DR PROSITE; PS00893; ABC TRANSPORTER 2; 2.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN 1.  
 KW ATP-binding.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 1459 AA; 163830 MW; 6A016A5A6E7AA95C CRC64;  
 Query Match 74.3%; Score 7444; DB 2; Length 1459;  
 Best Local Similarity 96.3%; Pred. No. 0;  
 Matches 1459; Conservative 0; Mismatches 0; Indels 56; Gaps 1;  
 QY 17 DNNVTWNTSNPDFTKCFQNTLVVWVPCFYLMACFPFYLYLSRHRDGYIOMTPLNKTKA 76  
 Db 1 DNNVTWNTSNPDFTKCFQNTLVVWVPCFYLMACFPFYLYLSRHRDGYIOMTPLNKTKA 60  
 QY 77 LGFLMWVCMADLFYSFWSRSGIFLAPVFLVSPITLLGTTLLATFLIOLERRKGVQSSG 136  
 Db 61 LGFLMWVCMADLFYSFWSRSGIFLAPVFLVSPITLLGTTLLATFLIOLERRKGVQSSG 120  
 QY 137 IMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDTFYVYFSLILLIQLVLSCFSDRSP 196  
 Db 121 IMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDTFYVYFSLILLIQLVLSCFSDRSP 180  
 QY 197 LFSETHDNPCCPSSASFLSRITFWITGLVIRGYRQPLEGSDMLSNKEDTSEQVVPV 256  
 Db 181 LFSETHDNPCCPSSASFLSRITFWITGLVIRGYRQPLEGSDMLSNKEDTSEQVVPV 240  
 QY 257 LVKNWKKCEAKTRKQPVKVYSSKDPAPQKESKVDANEVEALIVKSPKQKNPFLFKV 316  
 Db 241 LVKNWKKCEAKTRKQPVKVYSSKDPAPQKESKVDANEVEALIVKSPKQKNPFLFKV 300  
 QY 317 LVKTTGPPFLMSFFPKAHLDMFMFGSPQILKLLIKFVNDTKAPDMQGYFYTVLLFVTACL 376  
 Db 301 LVKTTGPPFLMSFFPKAHLDMFMFGSPQILKLLIKFVNDTKAPDMQGYFYTVLLFVTACL 360  
 QY 377 QTLVLHQYFPHICFVSGMRITKTAIVGAVYRKALVITNSARKSSTVGISVNLMSVDAQRFWD 436



Db 361 QTLVLHQYFHCISVGRMKTAVIGAVYRKALVITNSARKSSTVGEIVNLMSVDAQRFMD 420  
 QY 437 LATYINMIWSAPLOVILALYLALWNLGSPVSLAGVAVMLVPPVNAVMAKTKTYQVAHMK 496  
 Db 421 LATYINMIWSAPLOVILALYLALWNLGSPVSLAGVAVMLVPPVNAVMAKTKTYQVAHMK 480  
 QY 497 SKDNRIKLMNEILNGIKVLKYLAWELAFKDKVLAIROBELKVLKKSAYLSAVGTFTWVCT 556  
 Db 481 SKDNRIKLMNEILNGIKVLKYLAWELAFKDKVLAIROBELKVLKKSAYLSAVGTFTWVCT 540  
 QY 557 PELVALCTFAVYVTTDENNDLDAQAFVSLALFNILRPNLPLNIPVSIIVQASVSLKRL 616  
 Db 541 PELVALCTFAVYVTTDENNDLDAQAFVSLALFNILRPNLPLNIPVSIIVQASVSLKRL 600  
 QY 617 RIFLSHEELEPPSIERRPVKGGGNTSITVRNATFTWARSDDPTTLNGITFTSIPGALVAV 676  
 Db 601 RIFLSHEELEPPSIERRPVKGGGNTSITVRNATFTWARSDDPTTLNGITFTSIPGALVAV 660  
 QY 677 VQVCGKSSLLSALLAEMDKVEGVAIKGSVAVYVPPQAWIONDSLRENILFGCOLEBPY 736  
 Db 661 VQVCGKSSLLSALLAEMDKVEGVAIKGSVAVYVPPQAWIONDSLRENILFGCOLEBPY 720  
 QY 737 YRSVIOACALLPDLBILPSGDRTEICEKGVNLSGGQKQKVSLARAVYSNADIYLFDDPLS 796  
 Db 721 YRSVIOACALLPDLBILPSGDRTEICE ----- 747  
 QY 797 AYDAHVGKHI FENVIGPKGMLKNKTRILVTHSMSVLPQVDVLIIVMSGGKISEMGSYOELL 856  
 Db 748 -----KTRILVTHSMSVLPQVDVLIIVMSGGKISEMGSYOELL 784  
 QY 857 ARDGAFAEFLRYASTEOBQDAENGVTGVSFGKEAKQKMGMLVTDVSAGKOLQRLSS 916  
 Db 785 ARDGAFAEFLRYASTEOBQDAENGVTGVSFGKEAKQKMGMLVTDVSAGKOLQRLSS 844  
 QY 917 SSSYSGDISRHNHSTAELQAKAEKTEWKLMEADKAQTGVKLSVYWDYMKAIGLFTSF 976  
 Db 845 SSSYSGDISRHNHSTAELQAKAEKTEWKLMEADKAQTGVKLSVYWDYMKAIGLFTSF 904  
 QY 977 LSIFLFCMCHVSALASNYLSLWTDPIVNGTQHTKVLRSVYGALGISQGLAVFGYSA 1036  
 Db 905 LSIFLFCMCHVSALASNYLSLWTDPIVNGTQHTKVLRSVYGALGISQGLAVFGYSA 964  
 QY 1037 VSIGGILASRCLHVDLLHILSPMSFFERTSGNLVNRFSKELDTVDSMIPEVKMPWG 1096  
 Db 965 VSIGGILASRCLHVDLLHILSPMSFFERTSGNLVNRFSKELDTVDSMIPEVKMPWG 1024  
 QY 1097 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPVYSHFN 1156  
 Db 1025 SLFNVIGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPVYSHFN 1084  
 QY 1157 ETLGLSVIRAFEEQERFTHQSDLKVDENOKAYPSIVANRWLAVRLECVGNCIVLFAAL 1216  
 Db 1085 ETLGLSVIRAFEEQERFTHQSDLKVDENOKAYPSIVANRWLAVRLECVGNCIVLFAAL 1144  
 QY 1217 FAVISRHSLSAGLVGLSVSYQVTTYNLWVRMSSEMETNIVAVERLKEYSETEKEAPW 1276  
 Db 1145 FAVISRHSLSAGLVGLSVSYQVTTYNLWVRMSSEMETNIVAVERLKEYSETEKEAPW 1204  
 QY 1277 QIQTAPPSSWPQGVKVEFRNYCLARYREDLDFVLRHINVTNGEKGKVGIVGTGAKGSSL 1336  
 Db 1205 QIQTAPPSSWPQGVKVEFRNYCLARYREDLDFVLRHINVTNGEKGKVGIVGTGAKGSSL 1264  
 QY 1337 TIQGLFRINESAGEEIIIDGINAKTGLHDLRFKTIIPQDPVLFSGSLRMNLDPPSOYSD 1396  
 Db 1265 TIQGLFRINESAGEEIIIDGINAKTGLHDLRFKTIIPQDPVLFSGSLRMNLDPPSOYSD 1324  
 QY 1397 EEVWTSLELAHLKDFVSALPDKLDEHCAEGGNSVGOQLVCLARALRKTKILVLDEA 1456  
 Db 1325 EEVWTSLELAHLKDFVSALPDKLDEHCAEGGNSVGOQLVCLARALRKTKILVLDEA 1384  
 QY 1457 TAAVDLETDDLQSTIRTOFECTVLTIAHRLNTIMDYTRVIVLQKEIQEYGAESDLIQ 1516  
 Db 1385 TAAVDLETDDLQSTIRTOFECTVLTIAHRLNTIMDYTRVIVLQKEIQEYGAESDLIQ 1444

QY 1517 ORGLFYSMAXDAGLV 1531  
 Db 1445 ORGLFYSMAXDAGLV 1459  
 RESULT 6  
 OSUQA0  
 ID Q9UQA0 PRELIMINARY; PRT; 1456 AA.  
 AC Q9UQA0;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Multidrug resistance protein (Fragment).  
 GN Name=MRP;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98008927; PubMed=9344652; DOI=10.1006/geno.1997.4950;  
 RA Grant C.E., Kurz E.U., Cole S.P., Deeley R.G.;  
 RT "Analysis of the intron-exon organization of the human multidrug-  
 resistance protein gene (MRP) and alternative splicing of its mRNA.";  
 RL Genomics 45:368-378(1997).  
 CC 1- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AF022824; AAB83981.1; JOINED.  
 DR EMBL; AF022825; AAB83981.1; JOINED.  
 DR EMBL; AF022826; AAB83981.1; JOINED.  
 DR EMBL; AF022827; AAB83981.1; JOINED.  
 DR EMBL; AF022828; AAB83981.1; JOINED.  
 DR EMBL; AF022829; AAB83981.1; JOINED.  
 DR EMBL; AF022830; AAB83981.1; JOINED.  
 DR EMBL; AF022831; AAB83981.1; JOINED.  
 DR EMBL; AF022832; AAB83981.1; JOINED.  
 DR EMBL; AF022833; AAB83981.1; JOINED.  
 DR EMBL; AF022834; AAB83981.1; JOINED.  
 DR EMBL; AF022835; AAB83981.1; JOINED.  
 DR EMBL; AF022836; AAB83981.1; JOINED.  
 DR EMBL; AF022837; AAB83981.1; JOINED.  
 DR EMBL; AF022838; AAB83981.1; JOINED.  
 DR EMBL; AF022840; AAB83981.1; JOINED.  
 DR EMBL; AF022841; AAB83981.1; JOINED.  
 DR EMBL; AF022842; AAB83981.1; JOINED.  
 DR EMBL; AF022843; AAB83981.1; JOINED.  
 DR EMBL; AF022844; AAB83981.1; JOINED.  
 DR EMBL; AF022845; AAB83981.1; JOINED.  
 DR EMBL; AF022846; AAB83981.1; JOINED.  
 DR EMBL; AF022847; AAB83981.1; JOINED.  
 DR EMBL; AF022848; AAB83981.1; JOINED.  
 DR EMBL; AF022849; AAB83981.1; JOINED.  
 DR EMBL; AF022850; AAB83981.1; JOINED.  
 DR EMBL; AF022851; AAB83981.1; JOINED.  
 DR EMBL; AF022852; AAB83981.1; JOINED.  
 DR EMBL; AF022853; AAB83981.1; JOINED.  
 DR HSSP; P08716; 1M70.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m...; IEA.  
 DR GO; GO:0000165; F:nucleotide binding; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR011527; ABC membrane\_1.  
 DR InterPro; IPR011140; ABC TM transp...  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR005292; MRP assoc.  
 DR InterPro; IPR000719; Prot Kinase.  
 DR Pfam; PF00664; ABC membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transporter; 2.  
 DR SMART; SM00382; AAA; 2.



DR TIGR00957: MRP\_assoc\_pro; 1.  
 DR PROSITE; PS50929; ABC\_TMIP; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 KW ATP-binding.  
 FT NON\_TER  
 SQ SEQUENCE 1 1456 AA; 163231 MW; 8DE8AAB22BC481F2 CRC64;  
 Query Match 74.1%; Score 7419.5; DB 2; Length 1456;  
 Best Local Similarity 96.1%; Pred. No. 0;  
 Matches 1456; Conservative 0; Mismatches 0; Indels 59; Gaps 1;  
 QY 17 DNNVTNNSNDPFTKCFONTLVWPCFYLWACFPFYFLYLSRHDGVIQMTPLNKTITA 76  
 DB 1 DNNVTNNSNDPFTKCFONTLVWPCFYLWACFPFYFLYLSRHDGVIQMTPLNKTITA 60  
 QY 77 LGFLLIWVCWADLFYSFWSERSGIFLAPVFLVSPTLGLITTLATFLIQLERRKGVQSSG 136  
 DB 61 LGFLLIWVCWADLFYSFWSERSGIFLAPVFLVSPTLGLITTLATFLIQLERRKGVQSSG 120  
 QY 137 IMLTFWLVALCALAILRSKINTALKEDAOVDLFRDITFYVYFSLILLIQLVLSLCSFSDRSP 196  
 DB 121 IMLTFWLVALCALAILRSKINTALKEDAOVDLFRDITFYVYFSLILLIQLVLSLCSFSDRSP 180  
 QY 197 LFSETHDNPCESSASFLSRIITFWITGLIVRGVROPLEGSDLSLWLNKEDTSQVVPV 256  
 DB 181 LFSETHDNPCESSASFLSRIITFWITGLIVRGVROPLEGSDLSLWLNKEDTSQVVPV 240  
 QY 257 LVKNKKCECAKTRKQPVKVVYSSKDPAPKSSSKVDANEVEBALIVKSPQKWNPSLFKV 316  
 DB 241 LVKNKKCECAKTRKQPVKVVYSSKDPAPKSSSKVDANEVEBALIVKSPQKWNPSLFKV 300  
 QY 317 LYKTFGPYPLMFFPKAIDHLMFSGPQILKLLIKFVNDTKAPDWQGYFTVLLFTACL 376  
 DB 301 LYKTFGPYPLMFFPKAIDHLMFSGPQILKLLIKFVNDTKAPDWQGYFTVLLFTACL 360  
 QY 377 QTLVLHQPCHICFVSGMRITKAVIGAVYKALVITNSARKSTVGEIVNLMSVDQRMD 436  
 DB 361 QTLVLHQPCHICFVSGMRITKAVIGAVYKALVITNSARKSTVGEIVNLMSVDQRMD 420  
 QY 437 LATYINMIWSAPLOVILALYLWNLGPSVLGAVVMVLMVPVNAVMAKTKTYQVAHMK 496  
 DB 421 LATYINMIWSAPLOVILALYLWNLGPSVLGAVVMVLMVPVNAVMAKTKTYQVAHMK 480  
 QY 497 SKDNRIKLMNEILNGIKVLKAWELA FKQVLAIRQBELKVLKSAVLSAVGTFTWVCT 556  
 DB 481 SKDNRIKLMNEILNGIKVLKAWELA FKQVLAIRQBELKVLKSAVLSAVGTFTWVCT 540  
 QY 557 PELVALCTFAVVTIDENNILDAQTAFAVSLALFNILRFPNLIPMWISSIVQASVSLKRL 616  
 DB 541 PELVALCTFAVVTIDENNILDAQTAFAVSLALFNILRFPNLIPMWISSIVQASVSLKRL 600  
 QY 617 RIFLSHEELEPDSIERPVPKGGTNSITVRNATFTWARSDDPTTLNGITFTSIPGALNAV 676  
 DB 601 RIFLSHEELEPDSIERPVPKGGTNSITVRNATFTWARSDDPTTLNGITFTSIPGALNAV 660  
 QY 677 VQVGGCKGSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIQNDSLRNILEFGCQLEBPY 736  
 DB 661 VQVGGCKGSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIQNDSLRNILEFGCQLEBPY 688  
 QY 737 YRSVIQACALLPDLEILPSGDRTEIGEGKGNLGGQKORVSLARAVYSNADYLFDDPLS 796  
 DB 689 -----KGNLGGQKORVSLARAVYSNADYLFDDPLS 721  
 QY 797 AVDAHVGKHFENVIPKGMKNKTRILVTHSMYLPQVDVITVMSGGKISMGVQELL 856  
 DB 722 AVDAHVGKHFENVIPKGMKNKTRILVTHSMYLPQVDVITVMSGGKISMGVQELL 781  
 QY 857 ARDGAFAELRYASTEQDAENGVTGSGPGKEAKQWENGMLVTDGAGLQRLSS 916  
 DB 782 ARDGAFAELRYASTEQDAENGVTGSGPGKEAKQWENGMLVTDGAGLQRLSS 841

QY 917 SSSYSGDISRHNSHTAELOKAEAKKEETWKLMEADKAOTGOVKLSVYWDYMKALGLFISF 976  
 DB 842 SSSYSGDISRHNSHTAELOKAEAKKEETWKLMEADKAOTGOVKLSVYWDYMKALGLFISF 901  
 QY 977 LSIPLFCMNVHSALASNYWLSLWTDPIVNGTOBHTKVRLSVYGALGTSQGIAPVGYMA 1036  
 DB 902 LSIPLFCMNVHSALASNYWLSLWTDPIVNGTOBHTKVRLSVYGALGTSQGIAPVGYMA 961  
 QY 1037 VSIIGILASRCLHVDLHLSILRSWSPRPTSGNLVNRFSKELDTVDMSIPEVIMKPMFG 1096  
 DB 962 VSIIGILASRCLHVDLHLSILRSWSPRPTSGNLVNRFSKELDTVDMSIPEVIMKPMFG 1021  
 QY 1097 SLFNVIAGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPVYSHFN 1156  
 DB 1022 SLFNVIAGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPVYSHFN 1081  
 QY 1157 ETLLGVSVIRAFEPQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNCIVLFAAL 1216  
 DB 1082 ETLLGVSVIRAFEPQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNCIVLFAAL 1141  
 QY 1217 FAVISRHSLSAGLVCLSVSYSLQVTTYLNWLVRMSSEMETNIIVAVERLKEYSETEKAPW 1276  
 DB 1142 FAVISRHSLSAGLVCLSVSYSLQVTTYLNWLVRMSSEMETNIIVAVERLKEYSETEKAPW 1201  
 QY 1277 QIQTAPPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTINGGKGVIGVGTGAGKSSL 1336  
 DB 1202 QIQTAPPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTINGGKGVIGVGTGAGKSSL 1261  
 QY 1337 TLGLFRINESAGEHIIIDGINIAKGLHDLRFKTIIPQDPVLPFGSLRMNLDPSQVSD 1396  
 DB 1262 TLGLFRINESAGEHIIIDGINIAKGLHDLRFKTIIPQDPVLPFGSLRMNLDPSQVSD 1321  
 QY 1397 BEWVTSLELAHLKDFVSALPKLDHECAEGENLSVGORQLVCLARALLRKTILVLEA 1456  
 DB 1322 BEWVTSLELAHLKDFVSALPKLDHECAEGENLSVGORQLVCLARALLRKTILVLEA 1381  
 QY 1457 TAAVLETDLLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVILVDKGBIQEYGAPSDLLQ 1516  
 DB 1382 TAAVLETDLLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVILVDKGBIQEYGAPSDLLQ 1441  
 QY 1517 QRGLFYSMAKDAGLV 1531  
 DB 1442 QRGLFYSMAKDAGLV 1456  
 RESULT 7  
 Q6UR05 PRELIMINARY; PRT; 1531 AA.  
 AC Q6UR05;  
 DT 05-JUL-2004 (TREMELrel. 27, Created)  
 DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)  
 DE Multidrug resistance-associated protein 1.  
 GN Name=MRP1;  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RN [2]  
 SEQUENCE FROM N.A.  
 RX MEDLINE=22404678; PubMed=12516967;  
 RA Ma L., Pratt S.E., Cao J., Dantzig A.H., Moore R.E., Slapak C.A.;  
 RT "Identification and characterization of the canine multidrug  
 resistance-associated protein";  
 RL Mol. Cancer Ther. 1:1335-1342(2002).  
 RN [2]  
 SEQUENCE FROM N.A.  
 RA Ma L., Pratt S.E., Cao J., Dantzig A.H., Moore R.E., Slapak C.A.;  
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AY363728; AAQ3148.1; --  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.

DR GO:0042626; F:ATPase activity, coupled to transmembrane m. . . ; IEA.  
 DR GO:0000166; F:nucleotide binding; IEA.  
 DR GO:0005215; P:transporter activity; IEA.  
 DR GO:0008610; P:transport; IEA.  
 DR InterPro: IPR003593; AAA ATPase.  
 DR InterPro: IPR011527; ABC membrane 1.  
 DR InterPro: IPR001140; ABC\_TM transp.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001395; Aldo/ket\_red.  
 DR InterPro: IPR005292; MRP assoc.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR Pfam: PF00664; ABC membrane; 2.  
 DR Pfam: PF00005; ABC trans; 2.  
 DR ProDom: PD000006; ABC transporter; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR TIGRfam: TIGR00957; MRP assoc\_pro; 1.  
 DR PROSITE: PS00929; ABC TMIF; 2.  
 DR PROSITE: PS00211; ABC TRANSPORTER 1; 2.  
 DR PROSITE: PS00893; ABC TRANSPORTER 2; 2.  
 DR PROSITE: PS00063; ALDO-KETO REDUCTASE 3; UNKNOWN 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN 1.  
 KW ATP-binding.  
 SQ SEQUENCE 1531 AA; 171790 MW; 746361a71c6158bd CRC64;  
 Query Match 72.6%; Score 7272; DB 2; Length 1531;  
 Best Local Similarity 91.6%; Pred. No. 0;  
 Matches 1402; Conservative 70; Mismatches 59; Indels 0; Gaps 0;  
 QY 1 MALRGFCSDGSDPPLWDMNVNTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPFFYFLYLSRH 60  
 DB 1 MALRGFCSDGSDPPEWDMNVNTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPFFYFLYLSRH 60  
 QY 61 DRYGTQMTPLNKTALGFLWLVWVADLFSFWERSRGIFLAPVFLVSPILLGTTLLA 120  
 DB 61 DRYGTQMTPLNKTALGFLWLVWVADLFSFWERSRGIFLAPVFLVSPILLGTTLLA 120  
 QY 121 TFLIQLERRKGQSSGIMLTFWLVALVICALATLRSKIMTALKEDAQVDFLDITFYVFS 180  
 DB 121 TFLIQLERRKGQSSGIMLTFWLVALVICALATLRSKIMTALKEDAQVDFLDITFYVFS 180  
 QY 181 LLLIQLVLSGSDRPLSETHDNPCESSASFLSRITFWITGLIVRGYRPLEGSD 240  
 DB 181 LLLIQLVLSGSDRPLSETHDNPCESSASFLSRITFWITGLIVRGYRPLEGSD 240  
 QY 241 LWSLNKEDTSQVVPVLVKNMKCEAKTRKQPVVYSSKDPQAPKESKVDANEEVAL 300  
 DB 241 LWSLNKEDTSQVVPVLVKNMKCEAKTRKQPVVYSSKDPQAPKESKVDANEEVAL 300  
 QY 301 IVKSPQKEWNPFLFKVLYKTGPPYFLMSFFFKAIHDLMMFGSPQILKLLIKFVNDTKAPD 360  
 DB 301 IVKTPQKEREPSLFKVLVYKTGPPYFLMSFFFKALHDLMMFAGPELKLILLNFVNDKKAPD 360  
 QY 361 WQGYFYTVLLFVTAQLTLVHQYFHCIFVSGMRKTAIVGAVYRKALVITNSARKSTV 420  
 DB 361 WQGYLYTALLFICACLOTLVHLQYFHCIFVSGMRKTAIVGAVYRKALVITNSARKSTV 420  
 QY 421 GEIVNLSVDAQRFMDLATYINMTWSAPLOVILALYLWNLGSPVLGAVVAMVLMVFN 480  
 DB 421 GEIVNLSVDAQRFMDLATYINMTWSAPLOVILALYLWNLGSPVLGAVVAMVLMVFN 480  
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMEINLNGIKVLKLYAWELAFKDKVLAIRQBELKVLK 540  
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMEINLNGIKVLKLYAWELAFKDKVLAIRQBELKVLK 540  
 QY 541 KSAYLSAVGTTWCTPPLVALCTFAVVTVDENNILDQAQAFVSLAFNLRLPPLNLP 600  
 DB 541 KSAYLAAVGTWCTPPLVALCTFAVVTVDENNILDQAQAFVSLAFNLRLPPLNLP 600  
 QY 601 MVISIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSPPPT 660  
 DB 601 MVISIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWARSPPPT 660  
 QY 661 LNTGFTSPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAQWIND 720

DB 661 LNTGFTSPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAYVPOQAQWIND 720  
 QY 721 SLRENILFGCOLLEPPYRSVIQACALLPDLBILPSGDRTEIGKGVNLSSGQKQKRVSLAR 780  
 DB 721 SLRENILFGCOLLEPPYRSVIQACALLPDLBILPSGDRTEIGKGVNLSSGQKQKRVSLAR 780  
 QY 781 AVYNADIYLPDDPLSAVDHVGKHIPENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIV 840  
 DB 781 AVYCDSDIYLPDDPLSAVDHVGKHIPENVIGPKGMLKNKTRILVTHSMYSYLPQVDVIV 840  
 QY 841 MSGGKISEMGVQELLARDGAFELRTYASTEOBDAEENGVTGCVSPGKEAKOMENG 900  
 DB 841 MTGGKISEMGVQELLARDGAFELRTYASTEOBDAEENGVTGCVSPGKEAKOMENG 900  
 QY 901 LVTDGAGQLQRLSSSSSYSGDISRHHNSTAEILQKAEKKEETWKLMEADKAQTGOVKL 960  
 DB 901 LVTDVAGQLQRLSSSSSYSGDISRHHNSTAEILQKAEKKEETWKLMEADKAQTGOVKL 960  
 QY 961 SVYDYMKAIGLFTSFLIFLFCNHNVSALASNTWLSLWTDTPVNGTQHTKVLRSYVG 1020  
 DB 961 SVYDYMKAIGLFTSFLIFLFCNHNVSALASNTWLSLWTDTPVNGTQHTKVLRSYVG 1020  
 QY 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFESKEL 1080  
 DB 1021 ALGISQGIATVFGYSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPSGNLVNRFESKEL 1080  
 QY 1081 DTVDSDIPEVIMKPMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQFYVASSRQL 1140  
 DB 1081 DTVDSDIPEVIMKPMGSLFNIVIGACIVILLATPIAAIIIPPLGLIYFFVQFYVASSRQL 1140  
 QY 1141 KRLESVSRSPYSHFNELTLLGVSVIRAFESQERFIRQSDLVKVDENQKAYPSIVANRWLA 1200  
 DB 1141 KRLESVSRSPYSHFNELTLLGVSVIRAFESQERFIRQSDLVKVDENQKAYPSIVANRWLA 1200  
 QY 1201 VRLECVGNCIVLFAALFAVSRHSLSAGLSVSLQVTTLYLNLVSRMSSEMETNIVA 1260  
 DB 1201 VRLECVGNCIVLFAALFAVSRHSLSAGLSVSLQVTTLYLNLVSRMSSEMETNIVA 1260  
 QY 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNYCLRYREDLDFVLRLHINVTINGG 1320  
 DB 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNYCLRYREDLDFVLRLHINVTINGG 1320  
 QY 1321 EKVICVGTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 DB 1321 EKVICVGTGAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380  
 QY 1381 SGSLRMNLDPPSOYSDSEVWTSLELAHLKDFVSALPKDLHCECAEGENLSVGQRLVCL 1440  
 DB 1381 SGSLRMNLDPPSOYSDSEVWTSLELAHLKDFVSALPKDLHCECAEGENLSVGQRLVCL 1440  
 QY 1441 ARALLRKTILVDEATAAVDLETDLLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 DB 1441 ARALLRKTILVDEATAAVDLETDLLIQSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVL 1500  
 QY 1501 DKGEIQSVGAPSLDLOQRLGYFSMAKDAGLV 1531  
 DB 1501 DKGEIRECGPDLLOQRLGYFSMAKDAGLV 1531  
 RESULT 8  
 Q8HXQ5  
 ID Q8HXQ5 PRELIMINARY; PRT; 1530 AA.  
 AC Q8HXQ5;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Multidrug resistance protein 1.  
 GN Name=MRP1;  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.



DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)	Db	1	DWNTWNTSPDFTKCFQNTVLMVPCFYLMACFPFFFLYLSRDRGVIQMTPLNKTCTA	60
DE	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	QY	77	LGFLLIWVCMADLFYSFWSRSGIFLAPVFLVSPSTLLGTTTLATFLIQLERRKGVSQG	136
GN	Name=MRP;	Db	61	LGFLLIWVCMADLFYSFWSRSGIFLAPVFLVSPSTLLGTTTLATFLIQLERRKGVSQG	120
OS	Homo sapiens (Human).	QY	137	IMLTFMLVALVCAIILRSKIMTALKEADAQVDLFRDITFYVYFSLIATOLVLSCFSDRSP	196
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Db	121	IMLTFMLVALVCAIILRSKIMTALKEADAQVDLFRDITFYVYFSLIATOLVLSCFSDRSP	180
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	QY	197	LFSETIHDNPNCPBSSASFLSRITFWITGLIVRGYRQPLEGSLWSLNKEDTSEQVVPV	256
NCBI	TaxID=9606;	Db	181	LFSETIHDNPNCPBSSASFLSRITFWITGLIVRGYRQPLEGSLWSLNKEDTSEQVVPV	240
[1]	SEQUENCE FROM N.A.	QY	257	LVKWKEKCAKTRKQPVKVYVSSKDPAPKPESSKVDANEVEALIVKSPQKWNPSLFKV	316
RP	MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;	Db	241	LVKWKEKCAKTRKQPVKVYVSSKDPAPKPESSKVDANEVEALIVKSPQKWNPSLFKV	300
RX	Grant C.E., Kurz E.U., Cole S.P., Deeley R.G.;	QY	317	LYKTFGPYFLMSRPFKAHDLMMFSGPQILKLIKFNVDTKAPDWQGVYFTVLLFVTACL	376
RA	"Analysis of the intron-exon organization of the human multidrug-	Db	301	LYKTFGPYFLMSRPFKAHDLMMFSGPQILKLIKFNVDTKAPDWQGVYFTVLLFVTACL	360
RT	resistance protein gene (MRP) and alternative splicing of its mRNA.;"	QY	377	QTLVLHGYFHICFVSGMIRIKTAVIGAVYRKALVITNSARKSSTVGEIVNLSMDAQRFMD	436
RL	Genomics 45:368-378(1997).	Db	361	QTLVLHGYFHICFVSGMIRIKTAVIGAVYRKALVITNSARKSSTVGEIVNLSMDAQRFMD	420
CC	-!- SIMILARITY: Belongs to the ABC transporter family.	QY	437	LATYINMWSAPLOVILALYLLMLNLPSPVLGAVAVVLMVFNVAVMAMTKTYQVAHMK	496
DR	EMBL; AF022824; AAB83982.1; JOINED.	Db	421	LATYINMWSAPLOVILALYLLMLNLPSPVLGAVAVVLMVFNVAVMAMTKTYQVAHMK	480
DR	EMBL; AF022825; AAB83982.1; JOINED.	QY	497	SKDNRIKLMNEILNGIKVLKYLAWELAFKDKVLAIRQELKVLKKSAYLSAVGFTTWWCT	556
DR	EMBL; AF022826; AAB83982.1; JOINED.	Db	481	SKDNRIKLMNEILNGIKVLKYLAWELAFKDKVLAIRQELKVLKKSAYLSAVGFTTWWCT	540
DR	EMBL; AF022827; AAB83982.1; JOINED.	QY	557	PFLVALCTPAVYVITDENNILDQAQAPVSLALFNILFPLNLIIPMWISSIVQASVSLKRL	616
DR	EMBL; AF022828; AAB83982.1; JOINED.	Db	541	PFLVALCTPAVYVITDENNILDQAQAPVSLALFNILFPLNLIIPMWISSIVQASVSLKRL	600
DR	EMBL; AF022829; AAB83982.1; JOINED.	QY	617	RIFLSHEELEPDSIERPPVKDGGTNSITVRNATFTTWARSDPPTLNGITTSIPEGALVAV	676
DR	EMBL; AF022830; AAB83982.1; JOINED.	Db	601	RIFLSHEELEPDSIERPPVKDGGTNSITVRNATFTTWARSDPPTLNGITTSIPEGALVAV	660
DR	EMBL; AF022831; AAB83982.1; JOINED.	QY	677	VGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPPQAWIQNDSLRENILFPGQLEEPY	736
DR	EMBL; AF022832; AAB83982.1; JOINED.	Db	661	VGQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPPQAWIQNDSLRENILFPGQLEEPY	688
DR	EMBL; AF022833; AAB83982.1; JOINED.	QY	737	YRSVIOACALLPDLIELPSDGRTEIGEKNLSGGQKQVSLARAVYSNADIYLFDDPLS	796
DR	EMBL; AF022834; AAB83982.1; JOINED.	Db	689	YRSVIOACALLPDLIELPSDGRTEIGEKNLSGGQKQVSLARAVYSNADIYLFDDPLS	688
DR	EMBL; AF022835; AAB83982.1; JOINED.	QY	797	AVDAHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILVMSGGKISEMGSYOELL	856
DR	EMBL; AF022836; AAB83982.1; JOINED.	Db	689	AVDAHVGKHFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILVMSGGKISEMGSYOELL	725
DR	EMBL; AF022837; AAB83982.1; JOINED.	QY	857	ARDGAFAEFLRTYASTEQEQDAEENGVTGVSFGPKAKQKMGMLVTDSDAGKQLQRLSS	916
DR	EMBL; AF022838; AAB83982.1; JOINED.	Db	726	ARDGAFAEFLRTYASTEQEQDAEENGVTGVSFGPKAKQKMGMLVTDSDAGKQLQRLSS	785
DR	EMBL; AF022839; AAB83982.1; JOINED.	QY	917	SSYSYSGDISRRHNSTAELOKAEAKKEETWKLMEADKAQTCQVVKLSVYWDYMKAIGLFISF	976
DR	EMBL; AF022840; AAB83982.1; JOINED.	Db	786	SSYSYSGDISRRHNSTAELOKAEAKKEETWKLMEADKAQTCQVVKLSVYWDYMKAIGLFISF	845
DR	EMBL; AF022841; AAB83982.1; JOINED.	QY	977	LSIFLFWCNHVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYGALGISOGIAVFGYSMA	1036
DR	EMBL; AF022842; AAB83982.1; JOINED.	Db	846	LSIFLFWCNHVSALASNYWLSLWTDPIVNGTQEHKTVRLSVYGALGISOGIAVFGYSMA	905
DR	EMBL; AF022843; AAB83982.1; JOINED.	QY	1037	VSTGGILASRCLHVDLLHLSILRSPMSFFERTPSGNLVNRFSEKELDTVDSMIPVIKFMFG	1096
DR	EMBL; AF022844; AAB83982.1; JOINED.	Db	906	VSTGGILASRCLHVDLLHLSILRSPMSFFERTPSGNLVNRFSEKELDTVDSMIPVIKFMFG	965
DR	EMBL; AF022845; AAB83982.1; JOINED.	QY	1097	SLPNVIGACIVILLATPIAAIIPPLGLIYFFVQRFYVASSROLKELESVSRSPVYSHEN	1156
DR	EMBL; AF022846; AAB83982.1; JOINED.	Db	966	SLPNVIGACIVILLATPIAAIIPPLGLIYFFVQRFYVASSROLKELESVSRSPVYSHEN	1025

Query Match 70.9%; Score 7104.5; DB 2; Length 1400;  
Best Local Similarity 92.4%; Pred. No. 0;  
Matches 1400; Conservative 0; Mismatches 0; Indels 115; Gaps 1;  
17 DWNTWNTSPDFTKCFQNTVLMVPCFYLMACFPFFFLYLSRDRGVIQMTPLNKTCTA 76  
|||||

QY 1157 ETLLGVSVIRAFEEQERFIHQSDLVKVDENQKAYPSIVANRWLAVRLCVGNCIVLFAAL 1216  
 DB 1026 ETLLGVSVIRAFEEQERFIHQSDLVKVDENQKAYPSIVANRWLAVRLCVGNCIVLFAAL 1085  
 QY 1217 FAVISRHSLSAGLVGLSVYSYLVQVTTYNLWLRMSSEMETNIVAVRLKEYSETEKEAPW 1276  
 DB 1086 FAVISRHSLSAGLVGLSVYSYLVQVTTYNLWLRMSSEMETNIVAVRLKEYSETEKEAPW 1145  
 QY 1277 QIQTAPSSPQVGRVFRNVCRLRYREDLDFVLRHINVTGGBKVGIVGRTGAGKSSL 1336  
 DB 1146 QIQTAPSSPQVGRVFRNVCRLRYREDLDFVLRHINVTGGBKVGIVGRTGAGKSSL 1205  
 QY 1337 TGLGFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMNLDPPFSQYSD 1396  
 DB 1206 TGLGFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLFSGSLRMNLDPPFSQYSD 1265  
 QY 1397 BEVWTSLELAHLKDFVSALPKDLHCECAEGENLSVGQRLVCLARALLRTKILVLDEA 1456  
 DB 1266 BEVWTSLELAHLKDFVSALPKDLHCECAEGENLSVGQRLVCLARALLRTKILVLDEA 1325  
 QY 1457 TAAVDETDLLIOSTIRTFQEDCTVLTAAHRLNTIMDYTRVILDKGIEQVGPASDLLQ 1516  
 DB 1326 TAAVDETDLLIOSTIRTFQEDCTVLTAAHRLNTIMDYTRVILDKGIEQVGPASDLLQ 1385  
 QY 1517 ORGLFYSMAKDAGLV 1531  
 DB 1386 ORGLFYSMAKDAGLV 1400

RESULT 10  
 MRP1 MOUSE  
 ID MRP1 MOUSE STANDARD; PRT; 1528 AA.  
 AC O35379;  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Multidrug resistance-associated protein 1 (ATP-binding cassette, sub-family C, member 1).  
 DE family C, member 1).  
 GN Name=Abcc1; Synonyms=Abcc1a, Abcc1b, Mdrap, Mrp;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]\_TaxID=10090;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;  
 RX MEDLINE=96251691; PubMed=8649356;  
 RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.,  
 RA Deeley R.G.;  
 RT "Structure and expression of the messenger RNA encoding the murine  
 RT multidrug resistance protein, an ATP-binding cassette transporter.";  
 RL Mol. Pharmacol. 49:962-971(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perceca G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,

RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Vexard R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Iehii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki K., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Havaehizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 CC -!- FUNCTION: May participate directly in the active transport of  
 CC drugs into subcellular organelles or influence drug distribution  
 CC indirectly (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF022908; AAB80938.1; -;  
 DR EMBL; AK029896; BAC26654.1; -;  
 DR HSSP; P08716; IMT0.  
 DR MGD; MGI:102676; Abcc1.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR011527; ABC\_membrane\_1.  
 DR InterPro; IPR001140; ABC\_TM\_transp.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR Pfam; PF00664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAMs; TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE; PS50929; ABC\_TMIF; 2.  
 DR PROSITE; PS0211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 2.  
 KW ATP-binding; Glycoprotein; Repeat; Transmembrane; Transport.  
 FT DOMAIN 1 33  
 FT TRANSMEM 34 54  
 FT DOMAIN 55 74  
 FT TRANSMEM 75 95  
 FT DOMAIN 96 100  
 FT TRANSMEM 101 121  
 FT DOMAIN 122 133  
 FT TRANSMEM 134 154  
 FT DOMAIN 155 172  
 FT TRANSMEM 173 193  
 FT DOMAIN 194 317  
 FT TRANSMEM 318 338  
 FT DOMAIN 339 364  
 FT TRANSMEM 365 385  
 FT DOMAIN 386 441  
 FT TRANSMEM 442 462  
 FT DOMAIN 463 465  
 FT TRANSMEM 466 486  
 FT DOMAIN 487 548  
 FT TRANSMEM 549 569  
 FT DOMAIN 570 591  
 FT TRANSMEM 592 612  
 FT DOMAIN 613 963  
 FT TRANSMEM 964 984  
 FT DOMAIN 985 1022  
 FT TRANSMEM 1023 1043  
 FT DOMAIN 1044 1086  
 FT TRANSMEM 1087 1107

FT	DOMAIN	1108	1108	Extracellular (By similarity).	
FT	TRANSMEM	1109	1129	15 (By similarity).	
FT	DOMAIN	1130	1200	Cytoplasmic (By similarity).	
FT	TRANSMEM	1201	1221	16 (By similarity).	
FT	DOMAIN	1222	1233	Extracellular (By similarity).	
FT	TRANSMEM	1234	1244	17 (By similarity).	
FT	DOMAIN	1245	1528	Cytoplasmic (By similarity).	
FT	DOMAIN	644	868	ABC transporter 1.	
FT	DOMAIN	1290	1524	ABC transporter 2.	
FT	NP_BIND	678	685	ATP (potential).	
FT	NP_BIND	1324	1331	ATP (potential).	
FT	CARBOHYD	19	19	N-linked (GlcNAc...)	(Potential).
FT	CARBOHYD	1003	1003	N-linked (GlcNAc...)	(Potential).
SQ	SEQUENCE	1528	AA; 171183	MM; 68FD13667D61DBBB	CRC64;
Query Match					69.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity					88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;					
QY	1	MALRGCADGSDPLDNDVNTWNTSNPDFTKCFQNTLVWVPCFYLWACFPFFYLYLSRH	60		
Db	1	MALRGCADGSDPLDNDVNTWNTSNPDFTKCFQNTLVWVPCFYLWACFPFFYLYLSRH	60		
QY	61	DRGYIQMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPITLLGITLLA	120		
Db	61	DRGYIQMTPLNKTALGFLWIVCWADLFYSFWERSRGIFLAPVFLVSPITLLGITLLA	120		
QY	121	TFLIQLERRKGVOSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS	180		
Db	121	TFLIQLERRKGVOSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS	180		
QY	181	LLLIQLVLSCFSDRPLSETHDNPCEPSSASPLSRITFWITGLVIRGVROPLEGSD	240		
Db	181	LVLVQLVLSCFSDRPLSETHDNPCEPSSASPLSRITFWITGLVIRGVROPLEGSD	240		
QY	241	LWSLNKEDTSEQVVPVLVKNKKCEAKTRKQPKVYVS - SKDPAQPKSSKVDANEVEEA	299		
Db	241	LWSLNKEDTSEQVVPVLVKNKKCEAKTRKQPKVYVS - SKDPAQPKSSKVDANEVEEA	299		
QY	300	LIVSPQKERNPLFKVLYKTFGPFYFLMSFFKALHDLMMFSGPOILKLLIKFVNDRKAP	359		
Db	301	LIVSPHDKRPFSLFKVLYKTFGPFYFLMSFFKALHDLMMFAGPKILELIINFVNDREAP	360		
QY	360	DWQGYFYTVLLFVTAQTLVLHVFHICFVSGMRKITAIVGAVYRKALVITNSARKSST	419		
Db	361	DWQGYFYTVLLFVSAQTLVHVFHICFVSGMRKITAIVGAVYRKALVITNSARKSST	420		
QY	420	VGEIVNLSVDAQRPMDLATYNNIWSAPLOVILALYLLNLNGLSPVAGVAVMLMVPV	479		
Db	421	VGEIVNLSVDAQRPMDLATYNNIWSAPLOVILALYLLNLNGLSPVAGVAVMLMVPV	480		
QY	480	NAVMAKTKTYQVAHMKSKDNRIKLMEILNGIKVLKYALAEAFKDKVLAIRQELKVL	539		
Db	481	NAVMAKTKTYQVAHMKSKDNRIKLMEILNGIKVLKYALAEAFKDKVLAIRQELKVL	540		
QY	540	KKSAYLSAVGFTTWCPTPLFVALCTFAVYVITDENNILDATAFVSLALFNILFPLNIL	599		
Db	541	KKSAYLAAGFTTWCPTPLFVALCTFAVYVITDENNILDATAFVSLALFNILFPLNIL	600		
QY	600	PWISSIVQASVSLKRLRIFLSHELEPDSIERRPVKDGCGTNSITVRNATFTTWARSDPP	659		
Db	601	PWISSIVQASVSLKRLRIFLSHELEPDSIERRPVKDGCGTNSITVRNATFTTWARSDPP	659		
QY	660	TLNGITFSPREGALVAVVGQVCGCKSLLSALLAEMDKVEGHVAIKGSVAIVPOQAION	719		
Db	660	TLNGITFSPREGALVAVVGQVCGCKSLLSALLAEMDKVEGHVAIKGSVAIVPOQAION	719		
QY	720	DSLRENILFGQQLBEPYRSVIOCALLPDLLEILPSGDRTEIGKGNVLSGGQKQVSLA	779		
Db	720	DSLRENILFGHPLQENYKAWAEACALLPDLLEILPSGDRTEIGKGNVLSGGQKQVSLA	779		
QY	780	RAVYSNADIYLFDDPLSAVDHVGKHIFENVIGPKGMLKNKTRILVTHSMYSYLPQVDVII	839		

Db	780	RAVYSNADIYLFDDPLSAVDHVGKHIFENVIGPKGMLKNKTRILVTHSMYSYLPQVDVII	839		
QY	840	VMSGKISEMGSYQELLARDGAPAEFLRTYASTEOQDAENGVTGVSPPGKEAKQMENG	899		
Db	840	VMSGKISEMGSYQELLARDGAPAEFLRTYASTEOQDAENGVTGVSPPGKEAKQMENG	899		
QY	900	MLVTDGAKQORQLSSSSSYSGDISRHNSHTAELOKAEKEETWKLMEADKATQGVK	959		
Db	897	MLVTDVGVKHQRLHLSNSSSGDTSQHSIAELQKAGA - KEETWKLMEADKATQGVQ	955		
QY	960	LSVYDYMKAIGLFTISFLFMCNHSALASNYWLSLWTD - PIVANGTOBHTKVRISV	1018		
Db	956	LSVYDYMKAIGLFTISFLFMCNHSALASNYWLSLWTD - PIVANGTOBHTKVRISV	1015		
QY	1019	YGALISQGIADVGSMAVSIIGGILASRCLHVDLLHSILRSPMSFFERTPGNLNRRFSK	1078		
Db	1016	YGALGILQCAAFPGSMAVSIIGGIFASRLHLDLLNVLRSFMSFFERTPGNLNRRFSK	1075		
QY	1079	ELDTVDSMIPVIVKMGSLFNVIACIVILLATPIAAIIIPGLIYFFVQRFVYVASSR	1138		
Db	1076	ELDTVDSMIPVIVKMGSLFNVIACIVILLATPIAAIIIPGLIYFFVQRFVYVASSR	1135		
QY	1139	QLKRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRW	1198		
Db	1136	QLKRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYYPSIVANRW	1195		
QY	1199	LAVRLECVGNCIVLFAALFAVISRHSISAGLVGLSVSYSLQVTTYLYNWLVRMSSEMTNI	1258		
Db	1196	LAVRLECVGNCIVLFAALFAVISRHSISAGLVGLSVSYSLQVTTYLYNWLVRMSSEMTNI	1255		
QY	1259	VAVERLKEYSTEKAPWQIOETAPPSPQVQVEFERNYCLRYREDLDFVLRHINVTIN	1318		
Db	1256	VAVERLKEYSTEKAPWQIOETAPPSPQVQVEFERNYCLRYREDLDFVLRHINVTIN	1315		
QY	1319	GGEKGVIGRTGAGKSSITLGLFRINSAEGLIIGDINIAKIGLHDLRFKTIIPQDPV	1378		
Db	1316	GGEKGVIGRTGAGKSSITLGLFRINSAEGLIIGDINIAKIGLHDLRFKTIIPQDPV	1375		
QY	1379	LFSSGLRNMLDPPFSQYSDDEEVTSLLEHLKDFVSALPKLDHCEAGENLSVGQROLV	1438		
Db	1376	LFSSGLRNMLDPPFSQYSDDEEVTSLLEHLKDFVSALPKLDHCEAGENLSVGQROLV	1435		
QY	1439	CLARALLRKTKILVLDATAVLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVI	1498		
Db	1436	CLARALLRKTKILVLDATAVLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVI	1495		
QY	1499	VLDKGEIQEYGAPELDLQORGLFYSMADAGLV	1531		
Db	1496	VLDKGEVREGAPSELQORGLFYSMADAGLV	1528		
RESULT 11					
Q810E4					
ID	Q810E4	PRELIMINARY;	PRT;	1532	AA.
AC	Q810E4;				
DT	01-JUN-2003	(TrEMBLrel. 24, Created)			
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)			
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)			
DE	ATP-binding cassette protein C1.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Sprague-Dawley; TISSUE=Spleen;				
RA	Yabuuchi H., Takayanagi S., Ishikawa T.;				
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.				
CC	-1- SIMILARITY: Belongs to the ABC transporter family.				
DR	EMBL; AF487549; AA085437.1; -.				
DR	HSSP; P08716; IMT0.				
DR	GO; GO:0016021; C:integral to membrane; IEA.				
DR	GO; GO:0005524; F:ATP binding; IEA.				

DR GO: 0042626; F:ATPase activity, coupled to transmembrane m. . . ; IEA.  
 DR GO: 0000166; F:nucleotide binding; IEA.  
 DR GO: 0005215; F:transporter activity; IEA.  
 DR GO: 0006810; P:transport; IEA.  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR011527; ABC\_membrane\_1.  
 DR InterPro: IPR001440; ABC\_TM\_transp.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001395; Aldo/ket\_red.  
 DR InterPro: IPR005292; MRP\_assoc.  
 DR Pfam: PF00664; ABC\_membrane\_2.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR ProDom: PD000006; ABC\_transporter; 2.  
 DR SMART: SM00382; AAA; 2.  
 DR TIGRFAMs: TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE: PS00929; ABC\_TM1F; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE: PS00893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE: PS00063; ALDO\_KETO\_REDUCTASE\_3; UNKNOWN\_1.  
 DR ATP-binding.  
 DR KW ATP-binding.  
 DR SEQUENCE 1532 AA; 171491 MW; 286939PF63FA3F68 CRC64;  
 Query Match 69.2%; Score 6932.5; DB 2; Length 1532;  
 Best Local Similarity 87.3%; Pred. No 0;  
 Matches 1339; Conservative 99; Mismatches 92; Indels 3; Gaps 3;  
 QY 1 MALRGFCSADGSDPLWDMNVTWNTGNPDKFQNTLVWVPCFYLMACFPFYLYLSRH 60  
 DB 1 MALSFCSGSDGSDPLWDMNVTWNTGNPDKFQNTLVWVPCFYLMWCFYLYLSRH 60  
 QY 61 DRGYQMTPLNKTALGFLWIVCWADLFYFWSRSGIFLAPVFLVSPPTLLGITLLA 120  
 DB 61 DRGYQMTPLNKTALGFLWIVCWADLFYFWSRSGIFLAPVFLVSPPTLLGITLLA 120  
 QY 121 TFLIOLERRKGVSSGIMLTFLVALVCAALILRSKIMTALKEDAQVDLFRDITFVYFS 180  
 DB 121 TFLIOLERRKGVSSGIMLTFLVALVCAALILRSKIMTALKEDAQVDLFRDITFVYFS 180  
 QY 181 LLLIQLVLSGSDPLFSETVDRNPNCPPESSASFLSRITFWITGLIYRGYRQPLEGSD 240  
 DB 181 LVIQLVLSGSDPLFSETVDRNPNCPPESSASFLSRITFWITGLIYRGYRQPLEGSD 240  
 QY 241 LWSLNKEDTSEGVVPLVKNWKECAKTRKQPVVYS-SKDPAQKESKVDANEVEEA 299  
 DB 241 LWSLNKEDTSEGVVPLVKNWKECAKTRKQPVVYS-SKDPAQKESKVDANEVEEA 299  
 QY 300 LTVKSPQKWNPLFKVLYKTFPGYFLMSFPFKAHDLMMFSGPOILLIKLFVNDTKAP 359  
 DB 301 LTVKSHKDRDPSLFKVLKTFPGYFLMSFYLKALHDLMMFAGPELLEILINFVNDREAP 360  
 QY 360 DMQGYFYTVLLFVTAQLTLVHQYFHCIFVSGMRITKAVIGAVYRKALVITNSARKSST 419  
 DB 361 DMQGYLYTALLFVSACLQTLALHQYFHCIFVSGMRITKAVIGAVYRKALVITNSARKSST 420  
 QY 420 VGEIVNLMSVDAQRFMDLATYINMTWSAPLOVLLALYLWNLGSLVAGVAVMLMPVP 479  
 DB 421 VGEIVNLMSVDAQRFMDLATYINMTWSAPLOVTLALYFLWNLGSLVAGVAVMLMPVP 480  
 QY 480 NAVMAKTKTYOVAMHKSNDRI KLMEITLNGI KVLKLYAWELAFKDKVLATROBELKVL 539  
 DB 481 NAVMAKTKTYOVAMHKSNDRI KLMEITLNGI KVLKLYAWELAFQDKVMNIRQBELKVL 540  
 QY 540 KKSAYLSAVGTFTWCTPPLFVALCTFAVYVTTIDENNILDAQTAFLVSLALFNILRPLNLL 599  
 DB 541 KKSAYLSAVGTFTWCTPPLFVALCTFAVYVTTIDENNILDAQTAFLVSLALFNILRPLNLL 600  
 QY 600 PMWISSIVQASVSLKRLRIFLSHELEPDSIERRPVKOGGNTSITVRNATFTWARSDDP 659  
 DB 601 PMWISSIVQASVSLKRLRIFLSHELEPDSIERWSIKDGGGMSITVRNATFTWARSDDP 660  
 QY 660 TLNGITFTSEPEGALVAVGVQCGKSGSLLSALLAEMDKVEGHVAKGSAVYVPOQAWION 719  
 DB 661 TLNGITFTSEPEGALVAVGVQCGKSGSLLSALLAEMDKVEGHVAKGSAVYVPOQAWION 720

QY 720 DSLRENILFGCOLEBPYRSYVIOACALLPDLBILPSGDRTEICEKGVNLSGQKQKVSIA 779  
 DB 721 DSLRENILFGPLQEHCHYKAVMEACALLPDLBILPSGDLTEICEKGVNLSGQKQKVSIA 780  
 QY 780 RAVYSNADIYLPDDPLSAVDHVGKHI PENVIGPKGMLKNKTRILVTHSKMSYLPQVDVVI 839  
 DB 781 RAVYCNSDIYLLDDPLSAVDHVGKHI PEKVGPMGLLKNKTRILVTHGISYLPQVDVVI 840  
 QY 840 VMSGGKISEMGSYQELLARDGAFABFLRTYASTEQDQAEENGVTGVSFGKEAKOMENG 899  
 DB 841 VMSGGKISEMGSYQELLARDGAFABFLRTYASTEQDQAEENGVTGVSFGKEAKOMENG 900  
 QY 900 MLVTDAGKQLQRLSSSSSSSGDISRHNSSTAEIQAKAEKBTWKLMEADKAQGTQVVK 959  
 DB 901 ILVTDAGKQLQRLSSSSSSSGDISRHNSSTAEIQAKAEKBTWKLMEADKAQGTQVVK 959  
 QY 960 LSVYWDYKATGLFISLIFLPMCHVYSALASNTWLSLWTD-PLVNGTOEHTTKVRLSV 1018  
 DB 960 LSVYWDYKATGLFISLIFLPMCHVYSALASNTWLSLWTD-PLVNGTOEHTTKVRLSV 1019  
 QY 1019 YGALGISQICIAVFGYSMAVSIIGGILASRCLHVDLLHSILRSPSPFFERTPSGNLVRFSK 1078  
 DB 1020 YGALGILQGVAVFGYSMAVSIIGGIFASRLHLDLQNLVLRSPSPFFERTPSGNLVRFSK 1079  
 QY 1079 ELDTVDSMIPVVKFMGSLFNVI GACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSR 1138  
 DB 1080 ELDTVDSMIPVVKFMGSLFNVI GACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSR 1139  
 QY 1139 QLKRLSVSRSPVSHFNELTGLSVIRAFEOERFIHQSDLKVDENOKAVYPSIVANRW 1198  
 DB 1140 QLKRLSVSRSPVSHFNELTGLSVIRAFEOERFIHQSDLKVDENOKAVYPSIVANRW 1199  
 QY 1199 LAVRLECVGNCITLPAALPAVIRSHLSAGLVGVSYSLOVTVTLNMLVRMSSEMETNI 1258  
 DB 1200 LAVRLECVGNCITLPAALPAVIRSHLSAGLVGVSYSLOVTVTLNMLVRMSSEMETNI 1259  
 QY 1259 VAVRLEKYSYSETEKAPWQIOETAPPSWQVGRVFRNYCLRYREDLDFVLRIHNVITN 1318  
 DB 1260 VAVRLEKYSYSETEKAPWQIOETAPPSWQVGRVFRNYCLRYREDLDFVLRIHNVITN 1319  
 QY 1319 GGEKGVIGRTGAGKSSITGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPV 1378  
 DB 1320 GGEKGVIGRTGAGKSSITGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPV 1379  
 QY 1379 LPSGSLRNWLPFSSYDEEVTWLSLEHLKDFVSALPKLDPHECAEGENLSVGQRLV 1438  
 DB 1380 LPSGSLRNWLPFSSYDEEVTWLSLEHLKDFVSALPKLDPHECAEGENLSVGQRLV 1439  
 QY 1439 CLARALLRKTILVLDATAAVDLETDLLIOSTIRTFQEDCTVLTIAHRLNTIMDYTRVI 1498  
 DB 1440 CLARALLRKTILVLDATAAVDLETDLLIOSTIRTFQEDCTVLTIAHRLNTIMDYTRVI 1499  
 QY 1499 VLDKGEIEYGNAPSLLQORGLFYSMAXDAGLV 1531  
 DB 1500 VLDKGEIEYGNAPSLLQORGLFYSMAXDAGLV 1532  
 RESULT 12  
 Q8CG09 PRELIMINARY; PRT; 1532 AA.  
 AC Q8CG09  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Multidrug resistance-associated protein 1.  
 GN Name=Mp1;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.



RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
 RA Yang Z., Li C.S.W., Shen D.D., Ho R.J.Y.;  
 RL Submitted (Oct-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL; AY170916; AA086532.1; -;  
 DR HSP; P08716; IWO.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.  
 DR GO; GO:0001666; F:nucleotide binding; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR011527; ABC membrane 1.  
 DR InterPro; IPR011440; ABC TM transp.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR001395; Aldo/ket\_red.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR Pfam; PF00664; ABC membrane; 2.  
 DR Pfam; PF00005; ABC tran; 2.  
 DR ProDom; PD000006; ABC transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRfam; TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE; PS00929; ABC\_TMIF; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE; PS00063; ALDO\_KETO\_REDUCTASE\_3; UNKNOWN\_1.  
 KW ATP-binding.  
 SQ SEQUENCE 1532 AA; 171557 MW; C18F3554DD85732A CRC64;

Query Match 69.2%; Score 6932.5; DB 2; Length 1532;  
 Best Local Similarity 87.3%; Pred. No. 0;  
 Matches 1338; Conservative 100; Mismatches 92; Indels 3; Gaps 3;

QY 1 MALRGCSDGSDPLDWDNVTNTNPDPTKCFQNTLVWVPCFYLWACFPFYLYLSRH 60  
 DB 1 MALRGCSDGSDPLDWDNVTNTNPDPTKCFQNTLVWVPCFYLWACFPFYLYLSRH 60

QY 61 DRGYIQMTPLNKTKALGFLMIVCWADLFYSFWSRSGIFLAPVFLVSPITLLGTTLLA 120  
 DB 61 DRGYIQMTPLNKTKALGFLMIVCWADLFYSFWSRSGIFLAPVFLVSPITLLGTTLLA 120

QY 121 TFLIQLERRKGVOSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180  
 DB 121 TFLIQLERRKGVOSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYFS 180

QY 181 LILLIQLVLSFSDRPLFSETIHDNCPPESSASFLSRITFTWITGLIVRGVROPLEGSD 240  
 DB 181 LVFIQLVLSFSDRPLFSETIHDNCPPESSASFLSRITFTWITGLIVRGVROPLEGSD 240

QY 241 LWSLNKEDTSBQVPEVLVKNWKECAKTRKQPKVYVYS- SKDPAQPKESSKVDANEVEA 299  
 DB 241 LWSLNKEDTSBQVPEVLVKNWKECAKTRKQPKVYVYS- SKDPAQPKESSKVDANEVEA 299

QY 300 LIVKSPQKEMNPSLFKVLKTFGRFELMSFFKAIHDLMMSGQILKLLIKFVNDYKAP 359  
 DB 301 LIVKSHKDRDPSLFKVLKTFGRFELMSFFKAIHDLMMSGQILKLLIKFVNDYKAP 360

QY 360 DWQGYFTVLLFVTAQTLVHLVHFHICFVSGMRITKAVIGAVYRKALVITNSARKSST 419  
 DB 361 DWQGYFTVLLFVTAQTLVHLVHFHICFVSGMRITKAVIGAVYRKALVITNSARKSST 420

QY 420 VGEIVNLMSVDAQRFMDLATYINNIWSAPLOVIALYLLMLNGLSPVLAVVAVMLVMPV 479  
 DB 421 VGEIVNLMSVDAQRFMDLATYINNIWSAPLOVIALYLLMLNGLSPVLAVVAVMLVMPV 480

QY 480 NAVMAMTKTYQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROELKVL 539  
 DB 481 NAVMAMTKTYQVAHMKSKONRIKLMNEILNGIKVLKLYAWELAFKDKVNMROELKVL 540

QY 540 KKSAYLSAAGTFTVCTPFLVALCTFAVYVITDENNILDQAQTAQFVSLAFNILFRFPLN 599  
 DB 541 KKSAYLSAAGTFTVCTPFLVALCTFAVYVITDENNILDQAQTAQFVSLAFNILFRFPLN 600

QY 600 PMVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKCGGGTNSITVRNATFTWARSDDP 659  
 DB 601 PMVISSIVQASVSLKELRIFLSHEELEPDSIERRPVKCGGGTNSITVRNATFTWARSDDP 660

QY 660 TLANGITFSIPEGALVAVVQVCGGKSSLLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQN 719  
 DB 661 TLANGITFAIPDGLAVVQVCGGKSSLLSALLAEMDKVEGHVTLKGSVAYVPPQAWIQN 720

QY 720 DLSRENILFGCOLBEPYRSVIOACALLPDLILPSGDRTEIGEGVNLGGOKORVSLA 779  
 DB 721 DLSRENILFGPQLQHCYKAVWEACALLPDLILPSGDLTEIGEGVNLGGOKORVSLA 780

QY 780 RAVYNSADIYLPDDPLSAVDHVGKHIIFENVIGPKMLKNKTRILVTHSMYSLPQVDVII 839  
 DB 781 RAVYNSDIYLLDDPLSAVDHVGKHIIFEKVVGPMGLLKNKTRILVTHGISYLPQVDVII 840

QY 840 VMSGGKI SEMGSYOELLARDGAFAPFLTYASTEQEADENGVTGVSQPGKEAKOMENG 899  
 DB 841 VMSGGKI SEMGSYOELLARDGAFAPFLTYASTEQEADENGVTGVSQPGKEAKOMENG 900

QY 900 MLVTDGAKQLQRLSSSSSYSGDISRHNSSTAELOKAEAKETWKLMEADKATQGVK 959  
 DB 901 MLVTDGAKQLQRLSSSSSYSGDISRHNSSTAELOKAEAKETWKLMEADKATQGVK 959

QY 960 LSVYWDYKAI GLFISFLSIFLFCMNVHVSALASNYWLSLTDD-PIVNGTQEHKTVRLSV 1018  
 DB 960 LSVYWNMYKAI GLCISFLSIFLFCMNVHVSALASNYWLSLTDDPIVNGTQEHKTVRLSV 1019

QY 1019 YGALGISOGIAVFGYSMAVSI GGILASRLAHVDLHLSILRSPMFFERTSGNLVNRSK 1078  
 DB 1020 YGALGISOGIAVFGYSMAVSI GGILASRLAHVDLHLSILRSPMFFERTSGNLVNRSK 1079

QY 1079 ELDTVDSMIPVIXKFMGSLFNVI GACIVILLATPIAAIIIPPLGLIYFFVQRFVASSR 1138  
 DB 1080 ELDTVDSMIPVIXKFMGSLFNVI GACIVILLATPIAAIIIPPLGLIYFFVQRFVASSR 1139

QY 1139 QKRLSVSRSPVYSHNETLLGVSVIRAFEEORFTHOSDLKVDENQKAYPSIVANRW 1198  
 DB 1140 QKRLSVSRSPVYSHNETLLGVSVIRAFEEORFTHOSDLKVDENQKAYPSIVANRW 1199

QY 1199 LAVRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLWVMSSEMETNI 1258  
 DB 1200 LAVRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLWVMSSEMETNI 1259

QY 1259 VAVERLKEYSETEKAPWQIQTAPPSSWPQGVFRPNYCLRYREDLDFVLRHINVTIN 1318  
 DB 1260 VAVERLKEYSETEKAPWQIQTAPPSSWPQGVFRPNYCLRYREDLDFVLRHINVTIN 1319

QY 1319 GGEKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINAKTGLHDLRKITIIPODPV 1378  
 DB 1320 GGEKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINAKTGLHDLRKITIIPODPV 1379

QY 1379 LFGSLAMNLDPPFSQVSDSEEVWTSLELAHLKDFVSALPDKLDHECARGENLSVGQQLV 1438  
 DB 1380 LFGSLAMNLDPPFSQVSDSEEVWTSLELAHLKDFVSALPDKLDHECARGENLSVGQQLV 1439

QY 1439 CLARALLRKTKILVLDDEATAVDLETDLLIQSTIRQFQEDCTVLTIAHRLNTIMDYTRVI 1498  
 DB 1440 CLARALLRKTKILVLDDEATAVDLETDLLIQSTIRQFQEDCTVLTIAHRLNTIMDYTRVI 1499

QY 1499 VLDKGEIQEYGAPELQQLQGLFYSMAKQAGLV 1531  
 DB 1500 VLDKGEIQEYGAPELQQLQGLFYSMAKQAGLV 1532

RESULT 13  
 Q810G9  
 ID Q810G9 PRELIMINARY; PRT; 1523 AA.  
 AC Q810G9;  
 DT 01-JUN-2003 (T-EMBLrel. 24, Created)  
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)



DE ATP-binding cassette protein C1 variant A.  
GN Name=Abcc1;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_Taxid=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;  
RA Takayanagi S., Iehikawa T.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to the ABC transporter family.  
DR EMBL; AY174892; AA044983.1; -;  
DR HSSP; P08716; 1MTO.  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0005524; F: ATP binding; IEA.  
DR GO; GO:0042626; F: ATPase activity, coupled to transmembrane m. . .; IEA.  
DR GO; GO:0000166; F: nucleotide binding; IEA.  
DR GO; GO:0005215; F: transporter activity; IEA.  
DR GO; GO:0006810; F: transporter; IEA.  
DR InterPro; IPR003593; AAA ATPase.  
DR InterPro; IPR011527; ABC membrane 1.  
DR InterPro; IPR001140; ABC TM transp.  
DR InterPro; IPR003439; ABC transporter.  
DR InterPro; IPR001395; ABC transporter.  
DR InterPro; IPR005292; MRP assoc.  
DR Pfam; PF00664; ABC membrane; 2.  
DR Pfam; PF00005; ABC tran; 2.  
DR ProDom; PD000006; ABC transporter; 2.  
DR SMART; SM00382; AAA; 2.  
DR TIGRFAMs; TIGR00957; MRP assoc. pro; 1.  
DR PROSITE; PS00929; ABC TM1F; 2.  
DR PROSITE; PS00211; ABC TRANSPORTER 1; 2.  
DR PROSITE; PS00893; ABC TRANSPORTER 2; 2.  
DR PROSITE; PS00063; ALDOXETO\_REDUCTASE\_3; UNKNOWN\_1.  
KW ATP-binding.  
SQ  
Query Match 68.8%; Score 6892; DB 2; Length 1523;  
Best Local Similarity 87.0%; Pred. No. 0;  
Matches 1333; Conservative 98; Mismatches 90; Indels 12; Gaps 4;  
QY 1 MALRGFCSADGSDPLDWNVTWNTGNPDKFQNTLVWVPCFYLWACFPFYLYLSRH 60  
DB 1 MALRSFCSGSDGSDPLDWNVTWNTGNPDKFQNTLVWVPCFYLWACFPFYLYLSRH 60  
QY 61 DRGYQMTPLNTKTKALGFLIMVWADLFYFWSRSGIFLAPVFLVSPITLLGITLLA 120  
DB 61 DRGYQMTPLNKAKTALGFLIMVWADLFYFWSRSGIFLAPVFLVSPITLLGITLLA 120  
QY 121 TFLIOLERRKGVQSSGIMLTFLWALVLCALILRSKIMTALKEDAQVDLFDITFVYVS 180  
DB 121 TFLIOFERKRGVQSSGIMLTFLWALVLCALILRSKIIISALKKDAQVDFRDSAFVLYFT 180  
QY 181 LLLIQLVLCFSDRGLPSETIHDNCPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
DB 181 LVFIQLVLCFSDRGLPSETIHDNCPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240  
QY 241 LWSLNKEDTSEQVPLVQWKKCKACTKQPVVYS-SKDPAPKESKVDANEVEEA 299  
DB 241 LWSLNKEDTSEVVPVLVNNWKKCKVSRQVRIYVAPPKDTKPKGSSQLDVAEVEEA 300  
QY 300 LLVKSPQKWNPSLFVLYKTFPGPYFLMSFFFKALHDLMMFSGPQTLKLLIKFVNDTKAP 359  
DB 301 LLVKSHKDRDPSLFVLYKTFPGPYFLMSFLYKALHDLMMFAGPEILELIINFVNDREAP 360  
QY 360 DMQGYFYTVLLFVYACQLTLVHQYFHCIFVSGMRITKTAIVGAVYRKALVITNSARKST 419  
DB 361 DMQGYLYTALLFVSACQLTLVHQYFHCIFVSGMRITKTAIVGAVYRKALVITNSARKST 420  
QY 420 VGEIVNLSVDAQRFMDLATYINMIWSAPLOVTLALYLWNLGPSVLGAVVWLVMPV 479  
DB 421 VGEIVNLSVDAQRFMDLATYINMIWSAPLOVTLALYLWNLGPSVLGAVVWLVMPV 480

QY 480 NAVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVLAIROBELKYL 539  
DB 481 NAVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAWELAFKDKVMNIRBELKYL 540  
QY 540 KKSAYLSAVGPTTWCTPFLVALCTFAVYVTTDENNILDQAQTAFLVSLALFNILRPPLNL 599  
DB 541 KKSAYLSAVGPTTWCTPFLVALCTFAVYVTTDENNILDQAQTAFLVSLALFNILRPPLNL 600  
QY 600 PMVISSIVQASVSLKRLRIFLSHEELEDPSIRRRPVKGGGNSITVRNATFTWARSDDP 659  
DB 601 PMVISSIVQASVSLKRLRIFLSHEELEDPSIRRSIKGGGNSITVRNATFTWARSDDP 660  
QY 660 TLNGITFISPEGALVAVVQVCGCKSSLSALLAEMDKVEGHVAIKGSVAYVPPQAWIQN 719  
DB 661 TLNGITFIPDGLVAVVQVCGCKSSLSALLAEMDKVEGHVTLKGSVAYVPPQAWIQN 720  
QY 720 DSLRENILFGCOLLEPYRSVIOACALLPDLEILPSGDRTEIGEGVNLGGCKQKQVSLA 779  
DB 721 DSLRENILFGRPLQEHYKAVMEACALLPDLEILPSGDLTEIGEGVNLGGCKQKQVSLA 780  
QY 780 RAVYSNADILYLDPLSADVAHVGHIPENVLGPKGMLKNKTRILVTHSMSYLPQVDVIL 839  
DB 781 RAVYCNDSIYLLDDPLSADVAHVGHIFEKVGPMGLLKNKTRILVTHGISYLPQVDVIL 840  
QY 840 VMSGGKISEMGSYQELLARDGAFABFLRTYASTEQEQDAENGVTGVSFGPKGKQKQENG 899  
DB 841 VMSGGKISEMGSYQELLARDGAFABFLRTYASTEQEQDAENGVTGVSFGPKGKQKQENG 900  
QY 900 MLVTSAGKQQLORQLSSSSSYSGDISRHNSHTAELOKAEKKEETWKLMEADKAQGVK 959  
DB 901 ILVTDVAVGKPL-----HSVVTNQHSSTAELOKS-GVKEETWKLMEADKAQGVK 950  
QY 960 LSVYWDYMKAIGLFISFLPMCNHVSALASNTWLSLWTD-PVNGTQEHKTVKRLSV 1018  
DB 951 LSVYWNMYKAIGLCISFLSFLPCNHVSALASNTWLSLWTDPRVNGTQENRFLSV 1010  
QY 1019 YGALGISQGIIVFGYSMAVSIIGGILASRCLHVDLLHSILSRSPSPFPPTSGNLNRFSK 1078  
DB 1011 YGALGILQGVAVFGYSMAVSIIGGIFASRRLHLDLQNLVSRPSPFPPTSGNLNRFSK 1070  
QY 1079 ELDTVDSDMIPVVKPMGSLFNVIACIVILLATPIAIIIPPLGLIYFFVQRYFVASSR 1138  
DB 1071 ELDTVDSDMIPVVKPMGSLFNVIACIVILLATPIAIIIPPLGLIYFFVQRYFVASSR 1130  
QY 1139 QLKRLSVRSRSPVYSHFNELLGVSVIRAFEOERFIHOSDLKVDENQKAYPSIVANRW 1198  
DB 1131 QLKRLSVRSRSPVYSHFNELLGVSVIRAFEOERFIHOSDLKVDENQKAYPSIVANRW 1190  
QY 1199 LAVRLCEVGNICVLFAALFAVISRHSLSAGLVGLSVSYSLQITAYLVNMLVRMSSEMETNI 1258  
DB 1191 LAVRLCEVGNICVLFAALFAVISRHSLSAGLVGLSVSYSLQITAYLVNMLVRMSSEMETNI 1250  
QY 1259 VAVERLKEYSETEKAPQIQTAPSPSPQVGRVFRNYCLRYREDLDFVLRHINVTIN 1318  
DB 1251 VAVERLKEYSETEKAPQIQTAPSPSPQVGRVFRNYCLRYREDLDFVLKHINVTIE 1310  
QY 1319 GGEKGVIGVTRTGAGKSSLTGLFRINESAEGEIIIDGINAKIHLHDLFKFTIIPQDPV 1378  
DB 1311 GGEKGVIGVTRTGAGKSSLTGLFRINESAEGEIIIDGINAKIHLHDLFKFTIIPQDPV 1370  
QY 1379 LFSGSLRWMLDPPFSQYSDSEEVWMTSLAHLKDFVSALPDKLDHECAEGGENISVGQRLV 1438  
DB 1371 LFSGSLRWMLDPPFSQYSDSEEVWMTSLAHLKDFVSALPDKLDHECAEGGENISVGQRLV 1430  
QY 1439 CLARALLKTKILVDEATAVADLETDDLIQSTIRTFQEDCTVLTIAHLRNTIMDYTRVI 1498  
DB 1431 CLARALLKTKILVDEATAVADLETDDLIQSTIRTFQEDCTVLTIAHLRNTIMDYTRVI 1490  
QY 1499 VLDKGEIYEGAPSDLLQOGLFYSMKADAGLV 1531  
DB 1491 VLDKGEIYEGAPSDLLQOGLFYSMKADAGLV 1523

## RESULT 14

Q68CP7 PRELIMINARY; PRT; 1215 AA.  
 AC Q68CP7; 25-OCT-2004 (T-EMBLrel. 28, Created)  
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)  
 DE Hypothetical protein DKF2p781G125.  
 GN Name=DKF2p781G125;  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Amalgam;  
 RG The German cDNA Consortium;  
 RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Oeanger A.,  
 RA Fobo G., Han M., Wiemann S.;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CR749835; CAH18691.1; -  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR001140; ABC\_TM\_transp.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR InterPro; IPR005292; MRP\_assoc.  
 DR InterPro; IPR00719; Prot\_kinase.  
 DR Pfam; PF00664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR TIGRFAMs; TIGR00957; MRP\_assoc\_pro; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 KW ATP-binding; Hypothetical protein.  
 SQ SEQUENCE 1215 AA; 134926 MW; E490ADIP5268F72E CRC64;

Query Match 60.7%; Score 6075; DB 2; Length 1215;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1202; Conservative 0; Mismatches 3; Indels 10; Gaps 1;

QY 327 MSFFPKAHLDMFSGPQILKLIKFNVDNTPADWQGYFYVLLFVTAACLOTLLVHQYFH 386  
 Db 1 MSFFPKAHLDMFSGPQILKLIKFNVDNTPADWQGYFYVLLFVTAACLOTLLVHQYFH 60

QY 387 ICFVSGMIKTAIVGAVRKALVITNSARKSTVGEIVNLSVDAQRFMDLATYINMWS 446  
 Db 61 ICFVSGMIKTAIVGAVRKALVITNSARKSTVGEIVNLSVDAQRFMDLATYINMWS 120

QY 447 APLQVILALYLLNLGSPVLGAVVMVLPVNAVAMKTKTYQVAHMKSKDNRIKLMN 506  
 Db 121 APLQVILALYLLNLGSPVLGAVVMVLPVNAVAMKTKTYQVAHMKSKDNRIKLMN 180

QY 507 EILNGIKVLKLYANELAFKQVLAIRQBELKVLKKSAYLSAVGFTWVCTPFLVALCTFA 566  
 Db 181 EILNGIKVLKLYANELAFKQVLAIRQBELKVLKKSAYLSAVGFTWVCTPFLVALCTFA 240

QY 567 VYVTIDENNILDAQTAFLVSLALFNLPPLNLPWISSIVOASVSLKRLRIFLSHEELE 626  
 Db 241 VYVTIDENNILDAQTAFLVSLALFNLPPLNLPWISSIVOASVSLKRLRIFLSHEELE 300

QY 627 PDSITERRPVKGGTNSITVRNATFTWASDPPTLNGITFSPISPGALVAVVQVCGCKSS 686  
 Db 301 PDSITERRPVKGGTNSITVRNATFTWASDPPTLNGITFSPISPGALVAVVQVCGCKSS 360

QY 687 LLSALLAEMDKVEGHVAIKGSVAVVPOQAMTQNDLSRENILFGQLEEPYRSVIQACAL 746  
 Db 361 LLSALLAEMDKVEGHVAIKGSVAVVPOQAMTQNDLSRENILFGQLEEPYRSVIQACAL 420

QY 747 LPDLEIILPSGDRTEIGKGNVLSGGQKORVSLAPVYSNADIVLFDPLSAVDHVGKHI 806  
 Db 421 LPDLEIILPSGDRTEIGKGNVLSGGQKORVSLAPVYSNADIVLFDPLSAVDHVGKHI 480

QY 807 FENVIGPKGMLKNKTRILVTHSMSYLPQVDVIIIVMSGGKISEMGYSQELLARDGAFAPFL 866  
 Db 481 FENVIGPKGMLKNKTRILVTHSMSYLPQVDVIIIVMSGGKISEMGYSQELLARDGAFAPFL 540

QY 867 RYASTEQEQDABEN-----GVTGVSQPGKEAKOMENGLVTDTSAGKOLQRLSS 916  
 Db 541 RYASTEQEQDABENSTVMDDEEAGVTGVSQPGKEAKOMENGLVTDTSAGKOLQRLSS 600

QY 917 SSSYSGDISRHNSHTAEQKAEETKWEADKAOTGOVKLVSYVDYKKAIGLFTSF 976  
 Db 601 SSSYSGDISRHNSHTAEQKAEETKWEADKAOTGOVKLVSYVDYKKAIGLFTSF 660

QY 977 LSIFLFCMNVHSALASNYLSLWTDPIVNGTQEHTKVRLSVYGALGISQGIAGVGSWA 1036  
 Db 661 LSIFLFCMNVHSALASNYLSLWTDPIVNGTQEHTKVRLSVYGALGISQGIAGVGSWA 720

QY 1037 VSIGILASRCHVDLLHSILRSPMSFFERTPSGNLVRFSKELDTVDSMIPEVIKPMFG 1096  
 Db 721 VSIGILASRCHVDLLHSILRSPMSFFERTPSGNLVRFSKELDTVDSMIPEVIKPMFG 780

QY 1097 SLFNVGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPVSHFN 1156  
 Db 781 SLFNVGACIVILLATPIAAIIIPPLGLIYFFVQRFYVASSRQLKRLSVSRSPVSHFN 840

QY 1157 ETLGVSIVIRAFEEQERFIHOSDLKVDENQKAYYPSIVANRWLAVRLECVGNCIVLFAAL 1216  
 Db 841 ETLGVSIVIRAFEEQERFIHOSDLKVDENQKAYYPSIVANRWLAVRLECVGNCIVLFAAL 900

QY 1217 FAVISRHSLSAGLVGLSVSYSLQVTVYVNLVRMSSEMTNIVAVERLKEYSETEKEAPW 1276  
 Db 901 FAVISRHSLSAGLVGLSVSYSLQVTVYVNLVRMSSEMTNIVAVERLKEYSETEKEAPW 960

QY 1277 QIQTAPPSSWPQVGRVFEFRNYCLRYREDLDFVLRLHNVITNGGKVGIVGRTGAGKSSL 1336  
 Db 961 QIQTAPPSSWPQVGRVFEFRNYCLRYREDLDFVLRLHNVITNGGKVGIVGRTGAGKSSL 1020

QY 1337 TLGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLFSGLRMNLDPFQSYSD 1396  
 Db 1021 TLGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLFSGLRMNLDPFQSYSD 1080

QY 1397 EEVWTSLELAHLKDFVSALPKLDHECAEGENLSVGORQLVCLARALLRKTILVLDEA 1456  
 Db 1081 EEVWTSLELAHLKDFVSALPKLDHECAEGENLSVGORQLVCLARALLRKTILVLDEA 1140

QY 1457 TAAVDLETDDLIOSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1516  
 Db 1141 TAAVDLETDDLIOSTIRTFQEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1200

QY 1517 QRGLFYSMAKDAGLV 1531  
 Db 1201 QRGLFYSMAKDAGLV 1215

## RESULT 15

Q80ZK8 PRELIMINARY; PRT; 1519 AA.  
 ID Q80ZK8  
 AC Q80ZK8  
 DT 01-JUN-2003 (T-EMBLrel. 24, Created)  
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
 DE Similar to ATP-binding cassette, sub-family C (CFTR/MRP), member 3 (Fragment).  
 DE (Fragment).  
 GN Name=Abcc3;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N; TISSUE=Liver;  
 RA Strausberg R.;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

Query Match	45.1%	Score 4522;	DB 2;	Length 1519;
Best Local Similarity	56.2%;	Pred. No. 3.3e-270;		
Matches 865;	Conservative 269;	Mismatches 269;	Indels 38;	Gaps 8
QY	8	SADGSDPLWDMNVNTWNTSPDFTKCFONTVLVWVPCFVLWACPPFVLYLSRHRDGVOM	67	
DB	4	SEGLASKFWDNSLSYTTWTPDLTPCFQNSLAWPCYILWAAUPCYLYLRHQLGYIVL	63	
QY	68	TPLNKTKTALGELLWIVCWADLFYSFWERSRGIFLAPVFLVSPTLTGITTLATFLIQLE	127	
DB	64	SWLSRLKLTALGVLLWCVSWDLFYSFHGLIHGSSPAPVFFVTPLVGVTMLLATLLIQYE	123	
QY	128	RKRGVQSSIMLTFWLVALCALILRSKIMTALKEDAQVDLPRDITFYVYFSLLIQLV	187	
DB	124	RLRGVQSGVLIIFWLLCVICAIIPFRKLSALAEKILDPFRFTFYIFALVFCALI	183	
QY	188	LCGFSRSLPFGETIHDNPNCPSSASFLSRITFWWITGLIVRGYRQPLEGSDLSLNKE	247	
DB	184	LCSCFEKPLFSPENLDNPNCPASAGFSRLSFWWTFRLAIGYRRPLEDRDLNSLSE	243	
QY	248	DTSEQVVPVLVRNMKKCAKTRKQPKVYVYSSKDPQAPQKSSKVDANEVEALIVSPK	307	
DB	244	DCSHKVQVRLEAWKQ---QNAQSGSQATATAEPKIPGE-----DAVLLAKPRK	289	
QY	308	EWNPSLFKVLYKTPGYPFLMSFPFKAHDLMMFSGPOILKLLKLFVNDTKAPDQGYFTT	367	
DB	290	SKQPSFLRALVETFTSSLMSACFNLIQNLGFGVNPQLLSILIRFISDPTAPTWWGFLLA	349	
QY	368	VLLFTVACIQTALVHOYFHI CFSVGMRIKTAVICAVYRKALVITNSARKSSTGEIVNLM	427	
DB	350	GLMFUSSTWQTLIHQYTHCI FVMALRURTAII GVIYRKALVITNSVKRESTGEMVNLM	409	
QY	428	SVDAQRFMDLATYINNMISAPLOVILALYLLWMLGPSVLGAVVMVLVNVNVMAMKT	487	
DB	410	SVDAQRFMDVSPFINLLASAPLOVILAIYFLWQILGPSLAGAVVILVILPLINGAVSMKM	469	
QY	488	KTYQVAHMKSKNRIKLMNEILNGIKVLKYAWELAFKDKVLAIROBELKVLKKSAYLSA	547	
DB	470	KTYQVKQNFKDSRIKLMSEILNGIKVLKYAWEPSPLEQVKGTRQSELOLLRKQAYLQA	529	
QY	548	VGTEFTWCCTPLFVALCTFAVVYVTTDENNLDAQATFVSLAENLIRFPLNILPMWISSIV	607	
DB	530	ISTFTWICTPFLVTLITGVVYVYVDESNVLDAREKAFVSLSLFNLIKILPLNMQLPQLISGT	589	
QY	608	QASVSLKRLRIEFLSHEELPDSIERPPVKQGGGTNSITVRNATFTTWARDSDPPTINGITES	667	
DB	590	QASVSLKRIQDFLQNFQNDLPQCVERTISPG---YAITIHGFTTWAQDLPPTLHSLNIQ	646	
QY	668	IPEGALVAVGVGCGKSSLSALLAEKMDKVEGHVAIKGSVAVYVPOQAWITQNDLSRENIL	727	
DB	647	IPKGLAVAVGVPGCGKSSLSALLGEMKEKLEGVSVKGSVAVYVPOQAWITQNTCLOENVL	706	

Qy	1493	DYTRVILVDKGEIQEYGAPSDLLQQRLGFYSMAKDAGL	1530
		: : : : :	
		: : : : :	
		: : : : :	
		: : : : :	
Dp	1481	DYNRVLDKGVAEPDPSVNIIAAGSIFYGMAKDAGL	1518

**THIS PAGE BLANK (USPTO)**